



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 106297

TO: Phuong Bui
Location: 9a09 / 9e12
Tu sday, October 21, 2003
Art Unit: 1638
Phone: 305-1996
Serial Number: 09 / 868546

From: Jan Delaval
Location: Biotech-Chem Library
CM1-1E07
Phone: 308-4498

jan.delaval@uspto.gov

Search Notes

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 01:49:44 ; Search time 6721 Seconds
(without alignments)
11315.429 Million cell updates/sec

Title: US-09-868-546A-1
Perfect score: 1859
Sequence: 1 Gaaacactgacagacgca.....aataaatttcttcttc 1859

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_hgt:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_hgt_hum:*
- 31: em_hgt_inv:*
- 32: em_hgt_other:*
- 33: em_hgt_mus:*
- 34: em_hgt_pln:*
- 35: em_hgt_rod:*
- 36: em_hgt_man:*
- 37: em_hgt_vrt:*
- 38: em_sy:*
- 39: em_hgtgo_hum:*
- 40: em_hgtgo_mus:*
- 41: em_hgtgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	777.4	41.8	1820	8	AB022732	AB022732 Glycyrhri
2	769.4	41.4	1806	8	AB001379	AB001379 Glycyrhri
3	735.4	39.6	1668	8	CAR012581	AJ012581 Cicer ari
4	720.2	38.7	1500	8	AB025016	AB025016 Lotus jap
5	716.6	38.5	1781	8	CAR238439	AJ238439 Cicer ari
6	499.8	26.9	1711	8	CAR239051	AJ239051 Cicer ari
7	415.8	22.4	104321	8	AP006142	AP006142 Lotus jap
8	394.8	21.2	1693	8	BT002874	BT002874 Arabidops
9	394.4	21.2	1519	8	BT004449	BT004449 Arabidops
10	393.8	21.2	1488	6	AX506493	AX506493 Sequence
11	393.6	21.2	1561	8	AY113869	AY113869 Arabidops
12	393.6	21.2	1612	8	AY050849	AY050849 Arabidops
13	386.6	20.8	1494	6	AX507583	AX507583 Sequence
14	386.6	20.8	1494	6	AX590025	AX590025 Sequence
15	386.6	20.8	1494	6	AX651453	AX651453 Sequence
16	386.6	20.8	1525	8	BT000090	BT000090 Arabidops
17	386.6	20.8	1620	8	AY065192	AY065192 Arabidops
18	384.6	20.7	1656	8	AY039844	AY039844 Arabidops
19	384.6	20.7	1656	8	AY087256	AY087256 Arabidops
20	384	20.7	1652	8	D78607	D78607 Arabidops
21	383.8	20.6	1479	6	AX507477	AX507477 Sequence
22	383.2	20.6	1503	6	AX506778	AX506778 Sequence
23	383.2	20.6	1503	6	AX651658	AX651658 Sequence
24	381	20.5	1653	8	AY086486	AY086486 Arabidops
25	378.4	20.4	1503	6	AY074823	AY074823 Arabidops
26	377	20.3	1503	6	AX505366	AX505366 Sequence
27	377	20.3	1664	8	AY087977	AY087977 Arabidops
28	366.4	19.7	1508	6	BT000794	BT000794 Arabidops
29	361.8	19.5	1509	6	AX412388	AX412388 Sequence
30	361.8	19.5	1509	6	AX506140	AX506140 Sequence
31	361.8	19.5	1509	6	AX651784	AX651784 Sequence
32	361.8	19.5	1509	6	BT001045	BT001045 Arabidops
33	361.8	19.5	1587	8	AY093766	AY093766 Arabidops
34	360.2	19.4	1580	8	D78606	D78606 Arabidops
35	357	19.2	1507	8	AY096511	AY096511 Arabidops
36	357	19.2	1740	8	AY065209	AY065209 Arabidops
37	354	19.0	1578	8	AY084273	AY084273 Arabidops
38	351.4	18.9	1674	6	BD081607	BD081607 Plant fat
39	351.4	18.9	1674	8	HTCYP81C	AJ000477 Helianthu
40	350.6	18.9	1557	6	AX507264	AX507264 Sequence
41	350.6	18.9	1557	6	AX652005	AX652005 Sequence
42	349.2	18.8	1596	8	AY054214	AY054214 Arabidops
43	345	18.6	976	8	CAR249800	AJ249800 Cicer ari
44	344.8	18.5	1719	8	HTCYP81L	AJ000478 Helianthu
45	337.4	18.1	1632	6	AX412401	AX412401 Sequence

ALIGNMENTS

RESULT 1
AB022732

LOCUS
DEFINITION

AB022732
Glycyrhiza echinata CYP Ge-31 mRNA for cytochrome P450, complete cds,
1820 bp mRNA linear PLN 21-APR-2000

ACCESSION
VERSION

AB022732
AB022732.1 GI:4200043
cytochrome P450.

KEYWORDS
SOURCE

Glycyrhiza echinata
Glycyrhiza echinata

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrhiza.

REFERENCE

1 (sites)

AUTHORS Nakamura, K., Akashi, T., Aoki, T., Kawaguchi, K. and Ayabe, S.
TITLE Induction of isoflavonoid and retrochalcone branches of the flavonoid pathway in cultured Glycyrrhiza echinata cells treated with yeast extract
JOURNAL Biotechnol. Biochem. 63 (9), 1618-1620 (1999)
MEDLINE 20008263
PUBMED 10540749
REFERENCE 2 (bases 1 to 1820)
AUTHORS Akashi, T., Aoki, T. and Ayabe, S.
TITLE Direct Submission.
JOURNAL Submitted (23-JAN-1999) Shin-ichi Ayabe, Nihon University, Department of Applied Biological Science, Kameino 1866, Fujisawa, Kanagawa 252-8510, Japan [E-mail: ayabe@brs.nihon-u.ac.jp, Tel: 81-466-84-3703, Fax: 81-466-80-1141]
FEATURES Location/Qualifiers
 source
 1..1820
 /organism="Glycyrrhiza echinata"
 /mol_type="mRNA"
 /db_xref="taxon:46348"
 /cell_type="cultured cells"
 /clone_lib="lambda Zap II"
 /note="Cultivar maintained at Medicinal Plant Garden, School of Pharmaceutical Sciences, Kitasato University"
 1..1820
 /gene="CYP Ge-31"
 91..1590
 /gene="CYP Ge-31"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="BAA74465.1"
 /db_xref="GI:420004"
 /translation="MDILSLSYSVFVLFIFNIVIRARKFNLPFGPSPIIGN LHLKRLPHRTFKLGSEKYGHVILMFGRSLVVVSSASFOQCFTKNDVILNRPRF LSGKYIFVNTTLGSTSGEHWNLRIITDALVSLNHRINSFSGRDETQRLITLRA DSSNTPAEIELSVRFYDMFTNIMRISGRYGEDCDMSDLOEASQFRDMSELQ LSGANNKDTMPLRFLDFENLEKRLDKDISGTFDAFLRGLIQEHRAKERANTMIDHL LNQDSPEYTDQIIKGLAMLAGDSAVTLEWSMLNHPVLRVKKVDELDT HVGQRLVDESDPKLSYLNKVINETRLYTPAPLLPHSTSDCETGGYKVPQDITV LINAWIHRDPELMEATTFKPERFEKGELEKLIAFQMGRRACFGSLGAIIRASMTL ALLIQCPDWKLTNGDKIDMAERDGFITKLVLPLKMKRSRVPINKVFKQ"
 BASE COUNT 519 a 433 c 398 g 470 t
 ORIGIN
 Query Match 41.8%; Score 777.4; DB 8; Length 1820;
 Best Local Similarity 70.6%; Pred. No. 3.9e-165;
 Matches 1055; Conservative 0; Mismatches 431; Indels 9; Gaps 1;
 75 TCTCATAGCTGCTCTTTCTGCTGTTCTATCTCTCGGCGTGAAGTTTGTCTTCCAAAGCA 134
 110 TGTCTCTACTCGCTCTTCTATCTTCTGCTCTATTTTTCATATTCACACATTTGTGATCAGACAA 169
 135 GAAATTTGAGAAACATACACAGAGTCTCTCTCTCTTCCCATATAGGAAACCTTAACC 194
 170 GAAATTCACAAACCTCCACCGGTCCACCTTGCCTTCCCATATCGAAACCTCCACC 229
 195 TCCTCGAACAGCCAAATCCACCGTTTCTTCCACAGCATGTGAAACAGTACGCGCAACGTGG 254
 230 ACCTCAAAACGCCCTCCACGCGACCTTCAAGGGGTGTGCGAAAGTACGCGCATGTGA 289
 255 TTTCCCTCTGTTTCGGTTCAGCTCTGGCGGTGTGTCATCTCTCTCCACAGCATACCAAG 314
 290 TTTCCCTATAGTTTGGATCAGCCCTCGTGGTGTGATTTCTCTCGCATCTGAATTCACAC 349
 315 AATGCTTCCACAAACAGCAGTGTGCTTGGCCCAACCGGCTACTCTCTCTCGGAAAT 374
 350 AGTGTTTACCAAAACAGAGCTGTCTGGCAACAGAGCCCGCTTCTCTCGGAAAT 409
 375 ACATCTTACAAACACACCGTAGGCTCTGCTCCACGCGGAGACACTGGCGCAACC 434
 410 ACATCTTTTACAACTACACACCTTGGGGTCCACATCTCTATGCGGAGCACTGGGCAACC 469
 435 TCCGCGCATCACCGCCCTGGAGCTCTCTCCACGCGGCTTCCACTCTCTCTCTCCGAA 494

470 TCGTGTATATCCGCGCTCGAGCTCTCTTTCAAACACCGCATCAACAGCTTCTCCGGAA 529
 495 TCGGAGCCGACGAGACGAGCGTCTGATCGAGGGTGTGCTGGCCAAAGACTCCGAACG 554
 530 TCGAAGGACGAGACACAGAGGCTCATACAGGGCTTCCGATGACTCATCCACCAAC- 588
 555 AGGAAGAGTTTTCGCGAGTGGAGATTAGTTCGATGTTCAACGACTTAACTTACCAACA 614
 589 -----TTTGCAGAAATAGAACTCAGTTACAGGTTCTAGACATGACATTCACCAACA 640
 615 TAATGAGATGATATCGGGAAGAGGTTTACGGAGAGAGTGTAGATGAAGACGTTG 674
 641 TCATGAGATGATCTCCGGGAAGAGGTACTATGGGAGGACTGTGACATGTTCAGATCTTC 700
 675 AGGAAGCGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGCTA 734
 701 AAGAGCAAGCCAGTTTCAGGACATGTTGTTGAGTTGTTGCGGTCCTGAGGCTCAACA 760
 735 ACAAGGAGATCATCTTGCCTTCTCAGGTGTTTCGATTTTCAGATGTGGAGAGCGCT 794
 761 ACAAGACCGACTTCATGTCATGCTCAGGTTCTCGACTTTGAAAACCTGGAGAGAGGC 820
 795 TAAAGAGTATCATTAAGAGGTACGATTCATCTTGAATGAATCCTTCATGAGACCGTG 854
 821 TCAAGGATATCAGTGGCAAAACCGATGCTCTTCTGAGAGGACTCATCCAAGAGCACCGC 880
 855 CCAGCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACCTGCAAGAGACCCAGC 914
 881 CCAAGAGGACGTCGAAATACCATGATGATCATCTCTCATCTTCCAAGACTCACAGC 940
 915 CTCAGTACTACCTGACCAAAATCATCAAAGGCTTCTCTGCGCATGCTTTTGGTGGAA 974
 941 CTGAGTACTACCGATCAAATCATCAAAGGCTTCTCTGCGCATGCTCTTGGCGGAA 1000
 975 CTGACTCATCACTGGAGCTTTAGAGTGTCTATTAATTTTATTTGAATCAACCAGAG 1034
 1001 CCGACTCGTCTGCTGTAACCTTTGGAGTGTGCGATGCCAATTTGTTGAACATCCAGAG 1060
 1035 TGTGAAGAGGCAAGAGATGAATTCGACACTCAAGTGGGCAAGACCGCTTGTAAATG 1094
 1061 TATTGAAGAGGTAAAGGATGAATTGGATACACATGTTGGCCAAGATCGCTTGTGATG 1120
 1095 AGTCAGACCTTCCAAACCTTCCATATCTTAGGAAGATCATCTTGTGAGACACTTAGGTGT 1154
 1121 AGTCAGACCTTCCAAACCTCTCTTACCTTAAACCGTATCAATGAGACGCTTTCGGTGT 1180
 1155 ACCCCCGGCCCCAATCTTAATACCTCATGTGCTTCAAGAGATATTAATTTGAAGAT 1214
 1181 ACACTCGGCTCGTGTGTACTACCGCATCTGAGTGTGAGTGTGAGTGTGAGTGTG 1240
 1215 TCAATATCCCAAGACACAAATTTGTGATCATTAATGTTGGGCGCATGCGAGAGATCCTC 1274
 1241 ACAAGTTTCAACAGACACGATAGTATGATTAATGCTGGGCCATTCACAGGAGCCCTG 1300
 1275 AGTTGTGAATGATGCCATGCTTTAACTCAGAGGTTTGTGATGTGGAAGAGAGAGAGA 1334
 1301 AATTGTGAGTGAAGCCACGACTTCAAGCTCAGAGGTTTGAGAAAAGAGGTTGAGTTG 1360
 1335 AAAAGTTGTGATGATTTGGCATGGGAAGAGGCTTGGCCAGAGAACCCATGGCTATGC 1394
 1361 AAAAGTTAAATTGCTTTGGATGGGAAGAGGCGATGTCCCGGAGAGGCTTGGCTATTC 1420
 1395 AAAAGTGTGAGCTTTACTTTGGGATTTGTGATTCATGTTTGTGATGGAAACGAGTAAGT 1454
 1421 GGGCAATTAGCATGACTTTGGCGTTATGATTCATGCTTTGATTTGGAAGCTTACAAATG 1480
 1455 AGGAAAGCTTGTATGACAGAGAACAAATTGGATTCACCTTGTCAAGGTTAATTTCCATGG 1514
 1481 GTGATAAGATTGATGCGAGAGAGGAGTGAATTCACCTTGAACAAAATTTGGTTCCCTTAA 1540
 1515 AGGCCATGTGCAAGGCTCGCCCACTTGCCTCAATTAATTTGGAATTTAATTAATTAAT 1569
 1541 AGGCCATGTGTAATCACGTCACAGGTTTCAAGCAGTATCAACAGGTTTCAAGCAGTATTCAT 1595

RESULT 2
AB001379
LOCUS
DEFINITION
AB001379 1806 bp mRNA linear PLN 21-APR-2000
Glycyrrhiza echinata CYP81E1 mRNA for cytochrome P450, complete
cds, clone: CYP Ge-3.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycyrrhiza echinata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae;
Glycyrrhiza.
1 (sites)
REFERENCE
AUTHORS
Akaishi, T., Aoki, T., Takahashi, T., Kameya, N., Nakamura, I. and
Ayabe, S.
TITLE
Cloning of cytochrome P450 cDNAs from cultured Glycyrrhiza echinata
L. cells and their transcriptional activation by elicitor-treatment
Plant Sci. 126, 39-47 (1997)
2 (sites)
Akaishi, T., Aoki, T., Kameya, N., Nakamura, I. and Ayabe, S.
Two new cytochrome P450 cDNAs (Accession Nos. AB001379 and
AB001380) from elicitor-induced licorice (Glycyrrhiza echinata L.)
cells (PGR97-167)
Plant Physiol. 115, 1288 (1997)
3 (sites)
Akaishi, T., Aoki, T. and Ayabe, S.
CYP81E1, a cytochrome P450 cDNA of licorice (Glycyrrhiza echinata
L.), encodes isoflavone 2'-hydroxylase
Biochem. Biophys. Res. Commun. 251 (1), 67-70 (1998)
9909223
PUBMED
9790908
REFERENCE
AUTHORS
Ayabe, S.
TITLE
Direct Submission
Submitted (24-FEB-1997) Shin-ichi Ayabe, Nihon University,
Department of Applied Biological Science, Kameino 1866, Fujisawa,
Kanagawa 252-8510, Japan (E-mail: ayabe@brs.nihon-u.ac.jp,
Tel: 81-466-84-3703, Fax: 81-466-80-1141)
On Sep 27, 1997 this sequence version replaced gi:1845558.
D89431: submitted (20-Nov-1996).
FEATURES
Source
1. .1806
/organism="Glycyrrhiza echinata"
/mol_type="mRNA"
/db_xref="taxon:46348"
/clone="CYP Ge-3"
/cell_type="suspension-cultured cells"
/clone_lib="lambda Zap II"
1. .1806
/genes="CYP81E1"
69.1568
/genes="CYP81E1"
/codon_start=1
/product="cytochrome P450"
/protein_id="BAA2422.1"
/db_xref="GI:2443348"
/translation="MEILSLSVFYLALPFIENIVIRARFKNLPDPPSLPIIGN
LHKRLPHRTFGKSEKGVFLWFGRLVVVSSASEFOCCFTKNDVLANPRF
LSGKIFYNYITLGSYGEHNRNLTALDVLNRSINSPSGIRDETQRLITRLA
DSSNTFAEMELSLRYDMTFNNIMISGKRYGDCDSDLQASOFDMVSELQ
LSGANNKTFMELRLFDLENKRLKDISGKTDAFLRGLIEHRTKKERANTMIDHL
NLQDSOPVYDQIIKGLALAMLLAGTSSAVTLEWSNLLNHPVLKVKKRLDT
HYGQRLVDESILPKLTLYLKNVINETRLTPAPLLPHSTSDENICNGYKVPQITIV
LINAHWRDPELMTATFKPERFEKGELEKLFAGFGREACRPEGGLATRAISMTL
ALLIQCFDWKLLNGKIDLAERDFTLTKLVPLKMKCRSPRVINKFQ"

BASE COUNT
ORIGIN
515 a 428 c 390 g 473 t

Query Match

41.4%; Score 769.4; DB 8; Length 1806;

Best Local Similarity 70.2%; Pred. No. 2.5e-163;
Matches 1050; Conservative 0; Mismatches 436; Indels 9; Gaps 1;

QY 75 TCTCATACGCTGTCCTTTTCTCTGGTTCTATTCTTCGGGCTGAAGTTTGTTCCTTCCAAAGCA 134
Db 88 TGTCTACTCTCGTCTTCTATCTTGTCTCTATTCTTTCATATTCACATTGTGATCAGAGCA 147
QY 135 GAAATTTGAGAAACATACACACAGCTCCTCTCTCTTCCCATAATAGGAAACCTTAAAC 194
Db 148 GAAATTTCAAAAACCTCCACCGGCTTCCACCTTGGCTTCCCATAATCGGAAACCTCCAC 207
QY 195 TCCTCGAACAGCAATCCACCGTTTCTTCCACGATGTGAAAACAGTACGGCAACGTGG 254
Db 208 ACCTCAAAAGCCCTCCACCGCACCTTCAAGGGTGTTCGGAAGTACGGCCATGTGT 267
QY 255 TTTCCCTCTGTTGGTTCAGTCTGGCGGTTCATCTCTCTCCACAGCATACCAAG 314
Db 268 TTTCCCTATGTTGGATCAGCCTCTGCTGCTGTTCTCTGCATCTGATTTCCAAAC 327
QY 315 AATGCTTCAACCAACACGACGTTCCTTGGCCAAACGGCTACCTTCTCTCGGAAAT 374
Db 328 AGTGTTCACATAAAACGACGTCTGCTGGCAACAGGCCCCGCTTCTCTCGGAAAT 387
QY 375 ACATCTTCTACAAACACGACCGTAGCTCTCTCCACGGCGAGCACTGGCGCAAC 434
Db 388 ACATCTTTTCAACTACACCACTTGGGGTCCACATCTCTATGGCGAGCACTGGCGCAAC 447
QY 435 TCCGCGCATACACGCGCTGACGCTCTCCACGACGCGGTCCACTCTCTCTCGGAA 494
Db 448 TCCGCTGTATCACGCGCTCTGACGCTCTTCAACACCGCATCAACAGCTTCTCGGAA 507
QY 495 TCCGAGCGACGAGACGAAAGCTCTGTATGACAGAGTGTGCTGCGCCAAAGAACTCGAAG 554
Db 508 TCCGAAGGACGAGACACAAAGGCTCTAACGAGGCTTGGCGATGACTCGTCCACGAA- 566
QY 555 AGGAAGAGTTTGGCGAGTGGAGATTAGTTCGATGTTCAAGACTTAATCTTACAAACA 614
Db 567 -----TTTGGCCGAAATGGAATCTAGCTCTATGACATGATGCTTCAACAACA 618
QY 615 TAATGAGATGATATCGGGGAAGAGTTTACGGAGAGAGAGTGAAGATGAAGACGTG 674
Db 619 TCATGAGATGATCTCTGGGAAGAGATTATATGGGAGAGACTGTGACAGCTGATCTCC 678
QY 675 AGGAAGCGAGGAGTTTCAAGAGACTGTGACAGAAATTTGGAACTCATGGCTTGGCTA 734
Db 679 AAGAAGCAAGCCAGTTTCAAGGACATGGTGTCTGAGTTGCTACAGTTGTCTGGGCAACA 738
QY 735 ACAAGGAGATCATTTGCTTCTCAGTGGTTCGATTTTTCAGATGTGGAGAGCGCT 794
Db 739 ACAAGACTGACTTTCATGCCATTGCTCAGGTTCTCTGAAACCTTAGAGAGAGGC 798
QY 795 TAAAGAGTATCAGTAAGAGGTACGATTCATCTTGAATAAGATCTTTCATGAGAACCGTG 854
Db 799 TCAGGACATCAGTGGGCAAAACCGATGCTTCTTGAGAGGACTCATGAGAGCACCGCA 858
QY 855 CCAGCAATGACCGCCAGAAATTCATGATTCATCTCTCAAACTGCAAGAGACCCAGC 914
Db 859 CCAAGAGAGCGGTGCAAAATACCATGATGATCATCTCTCAATCTCCAAGACTCACAGC 918
QY 915 CTCAGTACTACATGACCAATCATCAAGGCTTGTCTGCGCATGCTTTTGTGGNA 974
Db 919 CTGAGTACTATACCGATCAAAATCATCAAGGCTTGTCTGCGCATGCTTCTTGGGGAA 978
QY 975 CTGACTCATCACTGGGACTTTAGAGTGGTCAATCTATCTAAATTTATTGAATCAACCCAGG 1034
Db 979 CCGACTCTCTGCTGTTACTTTGGAGTGTGCTGCTCAATTTGTTGAACCATCCAGAG 1038
QY 1035 TGTGGAAGAGCGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTGTAAATG 1094
Db 1039 TATTGAAGAGGTAAGGATGAATTTGGATACACATGTTGGCAAGATCGCTTGGTAGATG 1098
QY 1095 AGTCAGACTTCCAAAACCTTCCATATCTTAGAGAGATCATCTTGGAGACACTTAGGTTGT 1154

Db 1099 AGTCAGACTTCCAAAACCTCACTTACCTTAAACAGTTATCAATGAGACGCTTCGGTGT 1158
 Qy 1155 ACCCCCGGCCCCAATCTTAATACCTCATGTGCTTCAGAGATATATACAAATGAAGAT 1214
 Db 1159 ACATCCGGCTCCGTTGTTACTACCGCACTGCATTCAGATGATGCAATATGGAGAT 1218
 Qy 1215 TCAATATCCACGAGACACAATTTGATCATTAATGTTGGGATGCGAGAGATCCTC 1274
 Db 1219 ACAAGGTTCCACAGACACCATAGTATTAATGCTGCGGCCATTACAGGAGCCCTG 1278
 Qy 1275 AGTTGTGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGAATGCGAGGAGGAGA 1334
 Db 1279 AATTGTGACCTGAAGCCACGACTTCAGACCTGAGAGGTTTGAGAAAAGGGTGAGTTG 1338
 Qy 1335 AAAAGTTGGTAGCTTTGGCATGGGAAGAAGGGTTGCCAGAGAACCCATGCTATGC 1394
 Db 1339 AAAAGTTAATGCTTTGGGATGGGAAGAGGCAATGTCGCGAGAGGCTTGGCTATTC 1398
 Qy 1395 AAAGTGCAGCTTACTTTGGGATGTTGATTCATTAATGCTTGAATGCTGGAACGATAGTG 1454
 Db 1399 GGGCAATTAGCATGACTTTGGGTTAATGATTCATTAATGCTTGAATGGAAGCTTATAAATG 1458
 Qy 1455 AGGAAAGCTTGATGATGACAGACAACTTGCATCCTTCAGGTTAATTCATTCATTC 1514
 Db 1459 GTGATAGATTGACCTGCGAGAGGATGATTCACCTTGACAAAGTTGGTCCCTTAA 1518
 Qy 1515 AGGCCATGTGAAGGCTCGCCCACTTGCCACTTAAATTTGGAATTTAATTAAT 1569
 Db 1519 AGGCCATGTGAATCAGTCCAGTCATCAACAGGTTTCAAGCATTAATCAT 1573

RESULT 3
 CAR012581 1668 bp mRNA linear PLN 22-OCT-1999
 LOCUS Cicer arietinum mRNA for cytochrome P450.
 DEFINITION Cicer arietinum mRNA for cytochrome P450.
 ACCESSION AJ012581
 VERSION AJ012581.1 GI:3850629
 KEYWORDS cyp81E3 gene; cytochrome P450.
 SOURCE Cicer arietinum (chickpea)
 ORGANISM Cicer arietinum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;
 Cicer.
 1
 Overkamp,S. and Barz,W.
 Cloning of two Cicer arietinum L. cDNA's encoding Cytochrome P450s
 highly homologous to Isoflavone 2'-Hydroxylase from Licorice
 Plant Physiol. 120, 935-935 (1999)
 2 (bases 1 to 1668)
 Overkamp,S.
 Direct Submission
 Submitted (04-NOV-1998) Overkamp S., Institute for Plant
 Biochemistry and Biotechnology, University of Muenster,
 Hindenburgplatz 55, 48143, GERMANY
 Location/Qualifiers
 1..1668
 /organism="Cicer arietinum"
 /mol_type="mRNA"
 /cultiivar="ILC3279"
 /db_xref="taxon:3827"
 /note="heterotrophic cell suspension culture"
 1..1668
 /gene="cyp81E3"
 1..1497
 /gene="cyp81E3"
 /EC_number="1.14.14.1"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="CAA10067.1"
 /db_xref="GI:3850630"
 /db_xref="SPTREMBL:O92RW6"
 /translation="MGILSLCYSLFYLSFFFIIRLLFQSRKFNLPPLPPSLPIGN"

BASE COUNT 500 a 338 c 321 g 509 t
 ORIGIN
 Query Match 39.8%; Score 735.4; DB 8; Length 1668;
 Best Local Similarity 68.9%; Pred. No. 1.2e-155;
 Matches 1028; Conservative 0; Mismatches 456; Indels 9; Gaps 1;
 Qy 72 TGGTCTCATACGCTGCTTTTCTCGGTCTATTCCTCGCGGTGAAGTTTGTTCCTCAAA 131
 Db 17 TGTGTGCTACTCTCTTTTATCTTTCTTTTTCATCATATTAGGCTTTTGTTCAT 76
 Qy 132 GCAGAAAATTTGAGAAACATACACACAGGTCCTCTCTCTCCATAAATAGGAAACCTTA 191
 Db 77 CRAAGAAAATTCAAAATATCCACACAGGTCACCTTCTCTTCCATAAATGGCAACCTCC 136
 Qy 192 ACCTCTCGAACAGCAATCCACCGTTTCTTCAACGATGTGAAACAGTAGTCGGCAAG 251
 Db 137 ATCATCTCAAGCTCCCTCCACCGCACCTTTAAGGGACTCTCTAAACACTTATGGTGATA 196
 Qy 252 TGGTTCCTCTGTTTGGTTCGCTGCTGCGCTGCTCATCTCTCTCCTCAACAGCATACC 311
 Db 197 TCAATTCCTTTGGTGGTTCACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 256
 Qy 312 AAGATGCTTCAACAAACAGAGCTTGCCTTGGCCAAACCGGTACCTTCTCTCTCGGAA 371
 Db 257 AACATGCTTCACTAAAACAGATGTTGTCTAGCAAAATAGGCCTCGTTTCTCTCGAA 316
 Qy 372 AATACATCTTACAAACACACCGTAGGCTCTGCTCCCAACGCGAGCATCTGGGCA 431
 Db 317 AATACATCTTCTACAAATTAATACCTTAGGATCCACCTCTTACAGTGAACACTGGCGCA 376
 Qy 432 ACCTCGCGCATCACGCGCTGGAGCTCTCTCCACGAGCGCGCTCCACTCTCTCTCCG 491
 Db 377 ACCTTGTGTATCATCTTCTCTGATGTTCTTTCGACACACCGCATCAACACTTCTCTG 436
 Qy 492 GAATCCGAGCGAGAGCAAGCGTCTGATGAGAGGTTGGTGTGCTGGCCAAAGACTCGA 551
 Db 437 GAGTCCGAGGAGATGAGACTCAACGATTAATACGAAATTTGGCCGAGGATTCATCACTT 496
 Qy 552 ACAGGAGAGTTTGGCGAGTGGAGATTAGTTCGATGTTCAACGATTAACCTTACACACA 611
 Db 497 CC-----TTTGCTGAAGTAGAACTTAGTTTTCAGATTCTACGATATGACCTTCAACA 547
 Qy 612 ACATAATGAGATGATATCGGGAAGAGGTTTACCGAGAGGAGTGAAGTGAAGAACG 671
 Db 548 ACATATGCAATGATCCCTGGAAAGAGATACACGAGAGAGATTGACATGATGATC 607
 Qy 672 TTGAGGAAGCGAGGAGTTTCAAGAGACTGTGACAGAAATTTGGAACCTCATGGCTTGG 731
 Db 608 TTCAAGAGCAAGTCAATTCAGGATATGGTTACTGAACCTGCTGCAATTTCTGTGTCGA 667
 Qy 732 CTAAAGGAGGATACACTTGCCTTTCTCAGGTGGTTCGATTTTTCAGAAATGTCGAGAGC 791
 Db 668 ATAACAAAACCTGATTTTCATGCTTTTCTTAAATTTTGTGATTTTGAACCTTGGAGAAA 727
 Qy 792 GCTTTAAAGAGTATCAGTAAGAGGTAGATTTCCATCTTGAATAAGATCTCTTCATGAGAAC 851
 Db 728 GGGTTAAGATATTTGCTGACAAAACGACGCAATCTTTGAGAGGCTCTCTCAAGAACAC 787
 Qy 852 GTGCCAGCAATGACCCGAGAAATTCATGATCATCTCTCTCAAACTGCAAGAGACCC 911
 Db 788 GCAACAGAGGAAACGTACAAACACCATGATGACCATCTTCTTAAATTTGCAAGAAATCAC 847
 Qy 912 AGCTCAGTACTACTGACCAATCATCAAGGCTTCTCTGGCCATGCTTTTGGTG 971

```

Db      848 AACCTGAGTACTACATGATCAGATCATCAAGGCGCTTGCTTTGGCCATGCTCCTTGGCTG 907
QY      972 GAATGATCATCAACTGGGACTTTAGTGGTCAATATCAATTTAATTAATGAATCAACCCAG 1031
Db      908 GAACCGACTCATCTGCTGTAACCTTTAGAGTGGTCAATGCTAATTTGTTGAACCATCCAG 967
QY      1032 AGCTGTTGAAGAGGCAAGAGTAATGAGACTCAAGTGGGCAAGCCGCTTGTTAA 1091
Db      968 AGATATTGAAGAAGGTAAGAGGATGAATGAGATATCATGTAGCAAGATCGTTTGGTAG 1027
QY      1092 ATCAGTTCAGACCTTCACAACTTCATATCTTTAGGAGATCACTCTTGAGACCTTAGCT 1151
Db      1028 ATGATTCAGACCTTCACAACTTACTATCTTAAATGTCACTAGAGAGCTTCGGT 1087
QY      1152 TGTACCCCGCGGCCCAATTTCTAATACCTCATGTCTTCAGAGATATTAACAATGAAG 1211
Db      1088 TGTGTAATCCCGCTCCTTGTATTATACCACTCAACTTCAGATGATGATTTGGAG 1147
QY      1212 GATTCATATCCAGAGACACAATTTGATCAATTAATGTTGGGGCATGCGAGAGATC 1271
Db      1148 GATATAAAGTACCACAGACACCATATGATTAATGATGATGGGCCATTCACAGATC 1207
QY      1272 CTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGTTTGTGTTGGAAGAGG 1331
Db      1208 CTAAAGTCAGAGTGGAGCCCAAGCTTCAGCCGAGAGGTTTGACAAAAGGGGAGA 1267
QY      1332 AGAAAAGTTGGTAGCATTTGGCATGGGAAGAGGCTTGGCCAGGAGAACCCATGGCTA 1391
Db      1268 TTGAAAAGGTGATTGCAATTTGGCATGGGAAGAGAGTGCTCTCGAGAAGCTTTGGCTC 1327
QY      1392 TGCAAGTGTGAGCTTACTTTGGGATTTGATTTCAATGTTTGTGCTGGAACGAGTAA 1451
Db      1328 TTGAAACAATTAGCATGACTTTGGCAATTAATGTTTCAATGCTTTGATGGAAACGTACAA 1387
QY      1452 GTGAGGAAAGCTTGTATATGACAGAGAACAATTTGGATCACTTGTCAAGGTTAATTCAT 1511
Db      1388 GTGATGACAGATGATGATGGAAGAGAGATGATGATGTTTAAACAATGATTTCCAT 1447
QY      1512 TGGAGCCATGTGCAAGGCTGCGCCACTTGCCACTAAATTTGAATTTAATTA 1564
Db      1448 TAAAGGCCATGTGTAATACTCGTCCGGTGTGTCACCAAGGTTTTCAAGTAATTA 1500

RESULT 4
AB025016 1500 bp mRNA linear PLN 01-FEB-2001
LOCUS    Lotus japonicus mRNA for cytochrome P450, complete cds.
DEFINITION
AB025016
ACCESSION
VERSION  AB025016.1 GI:7415995
KEYWORDS  cytochrome P450.
SOURCE    Lotus japonicus
ORGANISM  Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
REFERENCE
1 (sites)
AUTHORS   Shimada,N., Akashi,T., Aoki,T. and Ayabe,S.
TITLE     Induction of isoflavonoid pathway in the model legume Lotus
japonicus: molecular characterization of enzymes involved in
Phytoalexin biosynthesis
JOURNAL   Plant Sci. 160 (1), 37-47 (2000)
PUBMED    11164575
REFERENCE 2 (bases 1 to 1500)
AUTHORS   Shimada,N., Aoki,T. and Ayabe,S.
TITLE     Direct Submission
JOURNAL   Submitted (16-MAR-1999) Shin-ichi Ayabe, Nihon University,
Department of Applied Biological Science; Kameino 1866, Fujisawa,
Kanagawa 252-8510, Japan (E-mail: ayabe@bri.nihon-u.ac.jp,
Tel:81-466-84-3703, Fax:81-466-80-1141)
FEATURES
Location/Qualifiers
1..1500
/organism="Lotus japonicus"

```

```

/mol_type="mRNA"
/db_xref="taxon:34305"
/clone="LjCYP-2"
/tissue_type="root"
/country="Japan: Gifu"
1..1500
/codon_start=1
/product="cytochrome P450"
/protein_id="BAA93634.1"
/db_xref="GI:7415996"
/translation="MDIISFLYSLFYVALFAIKLFLGSRKFKMLPQPTSLPIIGN
LHLKRPURHTFRALSEKYGDFVSLWFGNRLVSVVVSFADVOECFTKNDVLIANPRF
LSGYIFNYTITLSTSYGERHNLRTISLVSLSNHRINSFSPRRDRTILRLKLA
BDSKNFSEVELTSRFDFTNMIMISGKRYGDEDCDTELQEKASFRDQNTWELLO
LSGANNKADFWPILRLVDFEGLEKRVKGISGKTRDFRLGLLQEHQDKQRTANTWIDH
LTLQESPEVYVYDQILKGLALAMLAGTDSASVTLWMSNCNVLNYPVLEKIKAEID
THVQDRLVDESDDIPKLYLKNVINETLRLTPAPLLPHSASDDCTTGGYKVPDRTI
VLINAWLRDPQLWTEATTFKPERFDKGBLEKLPFLGLGRRAACFGELLALRAISMT
LALLIQCFDWRVSDDEIDMGERDGFVLMKSIPIVKAMCKSRPVNNVFK"
BASE COUNT      385 a      378 c      353 g      384 t
ORIGIN
Query Match      38.7%; Score 720.2; DB 8; Length 1500;
Best Local Similarity 68.8%; Pred. No. 3.3e-152;
Matches 1024; Conservative 0; Mismatches 453; Indels 12; Gaps 2;
QY      75 TCTCATAGCTGTCTCTTTCTCTGGTTCATCTCTCGCGGTGAAGTTTGTGTTTCCAAAGCA 134
Db      20 TTTACTATTTCTCTTCTACGTAGCTTTGTTGGCATATATCAACTCTTCTCTGGCTCA 79
QY      135 GAAATTTGAGAAACATACCAACAGGTCTCTCTCTCTTCCCATAATAGAGAAACCTTAACC 194
Db      80 GAAAGTTCAAAAACCTCTCTCCCGTCCAACTCACTTCCCATAATTTGCAACCTTCAACC 139
QY      195 TCCTCGAACAGCCATCCACCGTTTCTTCCAAAGCATGTCCGAAACAGTACCGCAAGTGG 254
Db      140 ACCTCAAAAGCGCCCTCCACCGCACCTTCAGGGCGCTGTCTGAGAAGTATGCGCAGCT 199
QY      255 TTTCCCTCTGGTTCTGGTTCAGGTCTGCGCGCTTGTCTCATCTCTCTCCAAACAGCATACCAAG 314
Db      200 TCTCCCTATGTTTCGAAACCGTCTTGTGGTGGTGTCTCTCTCTCGCGCATGTTCCAG 259
QY      315 AATGCTTTCACAAACACGAGTGGCTTTGGCCAAACCGGCTACCTTCTCTCTCGGAAAT 374
Db      260 AATGCTTTCACAAACACGAGTGGTGTCTTGGAAACCGGCTTCTCTCTCTCGGAAAT 319
QY      375 ACATCTTCTACAAACACGAGTGGTGTCTTCTCTCCGCGGAGCATGCGGCGCAACC 434
Db      320 ACATCTTCTACAACTACACCACTTACCTTACCTCTCTAGCGTGAGCATGCGCGTAAACC 379
QY      435 TCCGCGCATCACCGCCCTCGACGCTCTCTCCAGCAGCGGTCTCACCTCTCTCCGAA 494
Db      380 TCCGCGCATCACCTCTCTTGGATGCTCTCTCCAAACCGCATCAACAGCTTCTCCCCGA 439
QY      495 TCCGAGGCGAGCAGAGCAAGCGTCTGATGAGAGGTGGTGTGGCCCAAGAACTCGAAG 554
Db      440 TCCGCGCGATGAAACACACGCGCTGATTAGAGAGCTGGCTGAGGACTCTGCTAAGAAC- 498
QY      555 AGGAGAGATTTGCGCGAGTGGAGATAGTTAGATGTTCAACGACTTAATTAACAACA 614
Db      499 -----TTCTCCGAAGTGGAGCTCACTCAAGGTTCTTCGACATGACCTTCAACAACA 550
QY      615 TAATGAGGATGATATCGGGGAAGAGGTTTTTACGAGAGAGAGATGAGATGAAGAAGCTTG 674
Db      551 TCATGAGAGATGATCTCCGGGAAGAGGTACTATGAGAGAGATTTGACATGACCGAGCTTC 610
QY      675 AGGAAGCGAGGAGTTTCAGAGAGACTGTGACAGAGAAATTTGGAACTCATGGGCTTGGCTA 734
Db      611 AGGAGGCGAGCGAGTTTCAGGAGACATGGTGTGACCGAGTTGCTCAGTTGCTCCGGGGGAATA 670
QY      735 ACAAGGAGATCACTTGCCTTTCTCAGGTGGTTCGATTTTCAGAAATGTGAGAGAGCGCT 794
Db      671 ACAAGGCGGATTTTCATGCTTATTTCTCAGGCTTGTGATTTTGAAGGGGTTTGAAGAGGG 730

```


538 AGTTGGCTGAAGATTCACTCCTCC-----TTTGTCTGAAGTAGAACTTAGTTTCA 588
QY 588 TGTTCACGACTTAACCTTACACACATAATGAGGATGATATCGCGAAGAGCTTTTACG 647
Db 589 GATTCTATGATGACCTTCAACACATCATCGGAATGATCTCTGGAAGAGATACG 648
QY 648 GAGAGGAGTGAAGAGAGCTTGAAGAGCGGAGGAGTTCAGAGAGACTGTGACAG 707
Db 649 GAGAAGATTGTGATGATGATGATCTTCAAGAGCCAGTCAATTCAGGATATGTTACTG 708
QY 708 AATGTTTGAACATCATGGCTTGGCTTAAACAGGAGATCACTTGGCTTCTCAGTGGT 767
Db 709 AACTGTTGCAATTTGTCTGGTGCAATTAACAAACTGATTTTCATGCCCTTCTTAATTTT 768
QY 768 TCGATTTTCAAGATTGAGAGAGCGCTTAAAGAGTATCAGTAAGAGTACGATTCATCT 827
Db 769 TTGATTTTGAACCTTGGAGAAAGGTTTGAAGATATTTGCTGACAAACCGACGATCT 828
QY 828 TGAATAGATCCTTCATGAGAACCGTCCAGCAATGACCGCCAGAAATTCATGATCATC 887
Db 829 TGAGAGGACTCCTTCAAGAACCAACGCAACAGAGGAGAACGTACAAACACCATGATTGACC 888
QY 888 ATCTCTCAAACTGCAAGAGACCCAGCTCTCAGTACTACCTGACCAAAATCATCAAGGCC 947
Db 889 ATCTCTAAATTTGCAAGATCAACACTGAGTACTACCTGATCATGATCATCAAGGCC 948
QY 948 TTGCTCTGGCCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGATGGTCA 1007
Db 949 TTGCTTTGGCTATGCTCTTCTTGGTGAACCGACTCATCTGCTGTAACCTTTAGAGTGTCAA 1008
QY 1008 TATCTAATTTTGAATCACCCAGAGTGTGGAAGAGGCAAGAGATGAATTTGGACATC 1067
Db 1009 TGTCTAATTTTGAACCATCCAGGATTTGAAGAGGATTAAGAGATGAATTTGGATCTC 1068
QY 1068 AAGTGGACACAGACCGCTGTTAAATGAGTACAGTACAGCTTCCAAACTTCCATATCTTAGGA 1127
Db 1069 ATGTAGACAGATCGTTTGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
QY 1128 AGATCATCTTGGACACTTATGAGTGTGACCCCGGCGCCCAATTTCTAATACCTCATGTGT 1187
Db 1129 ATGTCTATGATGAACGCTTTCGGTGTGTGCTCTCTGCTCTCTCTCTCTCTCTCTCTCT 1188
QY 1188 CTTTCAAGATATTAACATTTGAAGATTTCAATATCCACAGACACAAATTTGATCATTA 1247
Db 1189 CTTTCAAGATATTTGAGGATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1248
QY 1248 ATGTTTGGGCGATGACAGAGATCTCAGTGTGGAATGATGACATGCTTTTAAACCTG 1307
Db 1249 ATGATGCGGCGATTCACAGGATCTTAAGTATGATGATGATGATGATGATGATGATGATGAT 1308
QY 1308 AGAGGTTTGTGTTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1367
Db 1309 AGAGGTTTGAACAAAGAGGAGATTCGAAAGGATGATGATGATGATGATGATGATGATGATGAT 1368
QY 1368 CTTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1427
Db 1369 TGTGCTCTGAGAGAGCTTTGCTCTCTGATCAATTTAGCATGATGATGATGATGATGATGATGAT 1428
QY 1428 AATGTTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1487
Db 1429 AATGTTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1488
QY 1488 TCACCTTTGTCAAGGTTAATTCATTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1547
Db 1489 TCGTCTTGACAAATTTGTTTCCATTAAGCCATGTTGTAACCTCGTCCGTCGTAAACA 1548
QY 1548 AATGGAATTTAATTAATTT 1566
Db 1549 AATTTTCAAGTAGCTAAT 1567

RESULT 6
CAR239051LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMCAR239051
Cicer arietinum mRNA for cytochrome P450 (cyp81E2 gene).
AJ239051
CYP81E2 gene; cytochrome P450.
Cicer arietinum (chickpea)
Cicer arietinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;
Cicer.REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL1
Overkamp, S., Hein, F. and Barz, W.
Cloning and characterization of eight cytochrome P450 cDNAs from
chickpea (Cicer arietinum L.) cell suspension cultures
Plant Sci. 155 (1), 101-108 (2000)
10773344
2 (bases 1 to 1711)
Overkamp, S.
Direct Submission
Submitted (17-MAY-1999) Overkamp S., Institute for Plant
Biochemistry and Biotechnology, University of Muenster,
Hindenburgplatz 55, Muenster, 48143, GERMANY
Location/Qualifiers
1. .1711
/organism="Cicer arietinum"
/mol_type="mRNA"
/cultivar="ILC 3279"
/db_xref="taxon:3827"
/country="Syria"
1. .1711
/gene="cyp81E2"
47. .1546
/gene="cyp81E2"
/EC_number="1.14.14.1"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAB43505.1"
/db_xref="GI:4894170"
/translation="MEYNFLTVIIFLSLFLLITLTKLFSKRKFKLPCCPPSPPIIGNL
HVKQPIHHTPVNLSKYGPITFLKGPVSVSSASLVEECFTKNDIIIFANRHS
KTKYLYNNNTWVSISYGLWRNLRRISSIELLSHRLNSFSEIKDETMLIGKLAK
KSKDPTKVERLSLFGELFTNTIMRVCKRFYGEESDGTNAEAKKFRDMNMEQF
GLGSLGDFVPLFLKDFSGHKLKVKGERMDALFQGLVDEHNNKNDKLNMTIDHL
LSQSPQDYSDQIIKGLINALIVAGTETSSVLEWMSNLNHPLEIKAKIELDN
LIGORHLIEAEATKLOYLQNIISITLRLHPVTMLPHLSHDCTIGDYVPRNTML
MYNAIHRDNLMDPMSKPERFENGQDGGFIPFGMRACPGSGALRLTLGLT
LGSLLQCFEWRKIGKEEVDMSGSGTVVPKALPLEAQCKARPIINKIFN"FEATURES
source547 a 311 c 342 g 511 t
BASE COUNT
ORIGINQuery Match 26.9%; Score 499.8; DB 8; Length 1711;
Best Local Similarity 60.7%; Pred. No. 2.7e-102;
Matches 892; Conservative 0; Mismatches 562; Indels 15; Gaps 4;
QY 99 TTCTATTCTCGCGCTGAAGTTTGTTCCTCAAGCAGAAAAATGAGAAACATACCAACCAG 158
Db 87 TTTTCTCATCTACAAATTTATTATTCTCCAAAAGAAAAATTCAAAAAACTTCCACCAT 146
QY 159 GTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 218
Db 147 GCCACCTTCGATTCCTATAATTTGGCACTCTCCATCAAGTTAAACAAACCAATTCACACA 206
QY 219 TCTTCAACGCATGTGAAACAGTAGCGCAACGTGGTTTCCCTCTCTCTCTCTCTCTCTCT 278
Db 207 CTTTCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAC 266
QY 279 TGGCGGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAC 338
Db 267 CGGTAGTAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAC 326
QY 339 CCTTGGCCAAACGGCTACCTTCTCTCTCGGAAAAATACATCTTCTACACACACACCG 398

Db 327 TTTTGGCAATCGTATTCACTCGCGCAAAACCAAGTATCTTGGCTACAAACACAAACA 386
Qy 399 TAGGCTCTGCTCCACGGCAGCACTGGCGCAACCTCCGCGCATCACCGCCTGACG 458
Db 387 TGGTTTCCATCGTACGGAGATCTTTGGCGCAACTTGGCGTGTATAGTTCAATTGAAA 446
Qy 459 TCCTCTCCACGAGCGGCTCACCTCTTCCGGAATCCGGAGCGAGCAGACGAGCGTC 518
Db 447 TACTCTCACTCACCGTCACTCTTTCGGAATCCGAAACGAGACGATGAGGT 506
Qy 519 TGATGAGAGGTTGGTGGCGCAAGAACTCGAACGAGGAAGATTTCGCGAGTGGAGA 578
Db 507 TGATCGGAAAA-----CTTGCTAAAAATCTCAC---AAAGATTTACAAAAGTTGAGC 557
Qy 579 TTAGTTCGATGTTCAAGCACTTAATCTACACACATAATGAGGATGATATCGGGGAAGA 638
Db 558 TTAGTTCCTTTTGGTGAGCTAACATTTAACTATATAGAAATGTTTGTGTAAC 617
Qy 639 GGTTTTACGAGAGAGAGTGAGATGAAGAACTGTGAGGAAGCGAGGAGTTTCAGAGAGA 698
Db 618 GATTTTATGGTGAAGATCCGATGGAACAAATGCTGAGGAGGCAAGAAGTTTAGAGACA 677
Qy 699 CTGTGACAGAAATGTTGAACTCATGCGCTTGGCTTAAACAGGAGATCATTGCTTTCC 758
Db 678 TGATCAATGAGATCAAGAGTTTGGGTTGGGTTCCAACTTGGTGAATTTGTGCTTTAT 737
Qy 759 TCAGTGTGTTCAATTTTTCAGAAATGTGAGAGAGCGTTAAAGAGATGATCAGTAAGAGTACG 818
Db 738 TCAAGTGTGTTGATTTTGTAGTGGTGTATAGAGAGTTTAAAGAGTTGGGAGAGAAATGG 797
Qy 819 ATTCATCTTGAATAAGATCCTTCATGAGAACCGTGGCAGCAATGACCGCA---GAATT 875
Db 798 ATGCATTTTCAAGGACTCGTTGACGAGCATCGTATACAAAGACAAAAATTTGATA 857
Qy 876 CCATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCTCAGTACTACACTGACCAAA 935
Db 858 CCATGATGATCATCTACTTTTCATGCAAGAAATCACAACCTGATTTATATAGTACCAAA 917
Qy 936 TCATCAAAAGGCTTGCTCTGCGCATGCTTTTGTGGAACCTGACTCATCACTGGGACTT 995
Db 918 TTATCAAGGGCTTATATAGCATGATGATGCTGGAACAGAACTTCATCTGTGACAT 977
Qy 996 TAGAGTGTCTAATCTAATTTATGATATCCAGAGAGTGTGGAAGAGCAAGAGATG 1055
Db 978 TAGAGTGGCCATGCTAATTTATTAACCATCCAGAAATTTGCAAAAAGCAAGATTG 1037
Qy 1056 AATTGGACACTCAAGTGGGACAGACCGCTTGTAAATGATGCTCAGACCTTCCAAACTTC 1115
Db 1038 AATTGGACAATCACATTTGCAAGAGCACTTAAATGAGGAAGCAAGCTACCAAAATTAC 1097
Qy 1116 CATATCTTAGGAAGATCATCTTGAGACACTTAGGTTGATCCCGCGGCCCAATTCTAA 1175
Db 1098 AATATCTTCAAAACATTAATCTCGAGACTTTGCGGTTGATCCACCGGTAACTATGTT 1157
Qy 1176 TACCTCATGTCTTCAGAGATATTAATGAAAGGATTCATATATCCAGAGACAA 1235
Db 1158 TGCCACATTTATCATCACACGATGTTATATAGGAGGCTATGATGTGCGACGTAACCA 1217
Qy 1236 TTGTGATCATTAATGTTGGGCGATGACAGAGATCCTCAGTTGTGGAATGATGCCACAT 1295
Db 1218 TGTGTATGGTGAATGATGGGCTATTCATAGGATCCAAATTTGTGGGCTGACCAATGA 1277
Qy 1296 GCTTTAAACCTCAGAGGTTTG---ATGTGGAAGAGAGGAGAAAAGTTGTGTAGCATTTG 1352
Db 1278 GCTTTAAGCCCGAGAGGTTTGAAGATGTTCAAGGGACATTTGTTGGCTTATTTCTTTG 1337
Qy 1353 GCATGGGAAGAGGCTTCCACAGAGAACCCATGGCTATGCAAGATGTTCAGCTTTACTT 1412
Db 1338 GGAATGGGAGGCGGCTTGTCTTGGGTCGGGCTAGCGCTTAAGAACAATTGGGCTTAACTT 1397
Qy 1413 TGGGATTTGATTCATTTGTTTGTGCTGGAACGAGTGAAGTGAAGAAAGCTTGAATGA 1472

Db 1398 TGGGCTCGTTGATTCATGCTTTAGTGGAAAAAGGATTGGGAAGGAAGTGTGACATGA 1457
Qy 1473 CAGAGAACAAATTTGGATCACCTTGTCAAGGTTAATTTCCATTGGAGGCCATGTGCAAGGCTC 1532
Db 1458 GTGAGGAAGCGGAACACTGTGGTGGCCAGGCCATTTCCATTAGAGGCCCATGCAAGAGCC 1517
Qy 1533 GCCCACTTCCCACTAAATTTGGAATTTAA 1561
Db 1518 GACCAATAATCAACGAAGATTTCAATTA 1546

RESULT 7
AP006142/c 104321 bp DNA linear PLN 19-MAR-2003
LOCUS Lotus japonicus genomic DNA, chromosome 2, clone:LJT08A24, TM0250,
DEFINITION complete sequence.
ACCESSION AP006142
VERSION AP006142.1 GI:29122779
KEYWORDS HTG.
SOURCE Lotus japonicus
ORGANISM Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
Lotus.
REFERENCE 1
AUTHORS Kaneko, T., Asamizu, E., Kato, T., Sato, S., Nakamura, Y. and Tabata, S.
TITLE Structural Analysis of a Lotus japonicus Genome. III. Sequence
Features and Mapping of Sixty-two TAC Clones Which Cover the 6.7 Mb
Regions of the Genome
JOURNAL DNA Res. 10, 27-33 (2003)
REFERENCE 2 (bases 1 to 104321)
AUTHORS Sato, S.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-2002) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 252-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337),
Fax: 81-438-52-3934)

FEATURES
1..104321
Location/Qualifiers
/organism="Lotus japonicus"
/mol_type="genomic DNA"
/db_xref="taxon:34305"
/chromosome="2"
/clone="LJT08A24"
/clone_lib="LJT library"
/note="TAC clone: TM0250"
BASE COUNT 31736 a 18441 c 19683 g 34461 t
ORIGIN

Query Match 22.4%; Score 415.8; DB 8; Length 104321;
Best Local Similarity 57.3%; Pred. No. 5.5e-83;
Matches 896; Conservative 0; Mismatches 567; Indels 100; Gaps 4;
Qy 78 CATACGCTGTCTTTTCTGTTCTATTCCTGGCGTGAAGTTGTTTTCACAAAGCAGAA 137
Db 8534 CCTACATCTCTGTTTCTTCAITTTTCTCATCACAGCCAAAGTTCTTACTCTCCAGAA 8475
Qy 138 AATTGAGAAACATACCACAGGTCCTCTCTTCCCATATAGGAACCTTTAACCTCC 197
Db 8474 GATTCAAGAACCTTCCACAGGCCACGCCCTTCTCCCTTATTCGGAACCTCCACAGC 8415
Qy 198 TCGAACAGCAATCCACCGTCTTCTTCCAAACGCGATGTGAAACAGTAGTCGGCAAGCTGTTT 257
Db 8414 TAAACCAACGATTCACACTTTTCCAGCCCTATCTCAGAAATATGTTCCCATTTTCT 8355
Qy 258 CCCTCTGTTGCGTTTCACTGCGGCTGTATCTCTCTCTCCAAAGCAGATACCAAGAT 317
Db 8354 CTTTCCGTTGGGTTCCCTACCGGTGCTGCTTTTCTCTGCTCCGCGCGGAAGAT 8295
Qy 318 GCTTACCAACACGAGCTTGGCTTGGCCACCGGCTACCTTCTCTCTCGGGAATAACA 377

Db 8294 GCTTACCAAAAAGCAGCATCATCTTCCGCAACCGTGTCCACTCCTCCAGACCAAGCACC 8235
QY 378 TCTTTACAAACACACACAGCTAGGCTCTCTGCTCCACGCGGAGCACTGGCGCAACCTCC 437
Db 8234 TCGCTACGACAGCAGCAGCAGCTCATCACCGCTTCTTACGCGGAGCACTGGCGCAACCTCC 8175
QY 438 GCGCATCACCGCCCTGGAGCTCTCTCCAGCAGCGGCTCCACTCTTCTCCGGAATCC 497
Db 8174 GCGGTATCAGCTCAATTTGAGATCTCTCAACCCACCGCCTCAACTCTCTTCTCCGAGTCC 8115
QY 498 GGAGCGACGAGACGAAGCGTCTGATGACAGAGTTGGTGTGGCCAGAGAACTCCAAACGAGG 557
Db 8114 GAAAGCAGAGACCATGAGGCTGCTCGGAGCTCGCTC-----ATPACAAAG 8067
QY 558 AAGAGTTTGGCGAGTGGAGATAGTTTCGATGTTCAACGACTTAACTTACAAACAAATAA 617
Db 8066 ATGAGTTTCGGAAGAAGTTGAGTCCAGTCCATGTTTGTGTAACCTAACTTCAACACCATCA 8007
QY 618 TGAGGATGATATCGGGAAGAGGTTTACGAGGAGAGTGAGATGAAGAGTTGAGG 677
Db 8006 TGAGATGGTGTGTGGCAACGCTACTACGCGGAGGAGTTTGACGGAACAGCGCCGAGG 7947
QY 678 AAGCGAGGAGTTTCAGAGAGACTGTGACAGAAATTTGGAACCTCATGGGCTTGGCTAAACA 737
Db 7946 AGCGAAGAAATTCAGAGATGTCATGATGAGATGCAACGTTTGGATTAGGCTTGGTC 7887
QY 738 AGGAGATACCTTGCCTTCTCAGTGGTTCGATTTTCAGATGTTGAGAGTGGAGAGCGCTTAA 797
Db 7886 TCGGTGAATTCGTGCTCTTCTTTCAGGTGTTGTTGATTTTCAGCGGAGAACCCAGCGGATTAC 7827
QY 798 AGAGTATCAGTAGAGGTAGATTCATTCATCTTGAATTAAGATCCTTCATGAGAACCGTGCCA 857
Db 7826 AGAACTTGGGGGAGAAAATGGAGCAGCTTTCAGAGGACTCATCGATGAATCGGAAATA 7767
QY 858 GCAATGACCCCGCAG---AATTCATGATCATCATCTCTCTCAAACTCGAAGACCCAGC 914
Db 7766 TGGGGGACCAAAATCAACACGATGATGATCATCTACTTCTTCCACAGAGACACAGC 7707
QY 915 CTGAGTACTACATGACCAAAATCAATCAAGGCTTGTCTGCG-----956
Db 7706 CTGAGTACTACATGACCAAAATTATCAAGGGCTTATTATGGTGGTGGGAAAATGTGAT 7647
QY 957 -----956
Db 7646 ATGTCATAAATGATTAAATAATAAATAATTATATTTATAATGTTGTTGTTTGTG 7587
QY 957 ----CCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGAGTGGTCAATATCT 1012
Db 7586 CAGTCGTTGATAGTGGCGGACAGAACTTCAGCTACACATTAGATGGGTGATGGCG 7527
QY 1013 AATTTATGTAATCACCCAGAGGTTTGAAGAGGCAAGAGATGAATTTGGACACTCAAGTG 1072
Db 7526 GAGCTGTTGAACCATCCGCGGTTTGTGGAGAAAGCGAGGAATGAGTTGGAGTCACAAAT 7467
QY 1073 GGCACAGACCGTGTAAATGAGTCAGACCTTCCAAACCTTCCATATCTTAGAGAGATC 1132
Db 7466 GGAATAGCGCGTAATGGAGGAGCGGAGGTAACTAAAGTGCCTTACTTTGGAGAAATA 7407
QY 1133 ATCTTTGAGACACTTAGTTGTACCCCGCGCCCAATTCATATCTCATGTGTCTTCA 1192
Db 7406 ATCTGAGACTTTGCGTTGATCCAGCAGTTCGATGTTGTATCCACATTTACTCTCT 7347
QY 1193 GAAGATATTAATTAAGGATTAATATATATCCACGAGACAAATGTTGATCATTAATGTT 1252
Db 7346 GAGGATTTGACGGTGGTGGCTATCAGTCGCGGTAATCAATATGTTGATGTTGATGCG 7287
QY 1253 TGGGGATGAGAGATCTCAGTTGTGGAATGATGCCACATGCTTTAACTTAACTGAGAG 1312
Db 7286 TGGGCTATACAGAGGACCCGAATTTGTGGTTGTATCCGTTGCAATTTGTCCTGAGAG 7227
QY 1313 TTGTG---ATGTGAAGAGAGAGGAGAAAAGTTGTTAGCATTTTGGCATGGGAAGAGGCT 1369
Db 7226 TTGAGANAATGTTGGTGGGATGCGCTGTTGGTCAATCTCTTTGGATGGGAGAGGTCA 7167

FEATURES

QY 1370 TCCCCAGGAGAACCCATGGCTATGCAAAAGTGTGAGCTTTACTTTGGGATTTGTTGATCAA 1429
Db 7166 TGTCTCGGTGGGCTTGGCCAGAGAACTATGGGTCTGACTCTGGGCTCGTTGATTCAG 7107
QY 1430 TGTTTTGTGCTGAAAAGAGTAAAGTGAAGAAAGCTTGATATGACAGAGAACTTTGGATC 1489
Db 7106 TGTCTTGTGATGGAAGAGATTGGCAGCGAAGAGATTGACATGACTTGAAGACGTGGAAT 7047
QY 1490 ACCTTGTCAAGGTAAATTCATTGAGGCGCATGTGCAAGGCTCGCCACTTGCACATAA 1549
Db 7046 ATCATACCTAAGCTATTCCTTTGAGGCGCTTTGCAAGCTCTGTCATCATCAACAG 6987
QY 1550 ATT 1552
Db 6986 ATT 6984

RESULT 8

BT002874
LOCUS
DEFINITION
Arabidopsis thaliana clone RAFL15-01-B10 (R20394) putative
cytochrome p450 family protein (At4g37320) mRNA, complete cds.

ACCESSION
BT002874.1 GI:27754486
VERSION
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana

REFERENCE
AUTHORS
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.

TITLE
Arabidopsis Full Length cDNA Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1693)
AUTHORS
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.

TITLE
Arabidopsis Full Length cDNA Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1693)
AUTHORS
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.

TITLE
Arabidopsis Full Length cDNA Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1693)
AUTHORS
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.

TITLE
Arabidopsis Full Length cDNA Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1693)
AUTHORS
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.

TITLE
Arabidopsis Full Length cDNA Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1693)
AUTHORS
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.

TITLE
Arabidopsis Full Length cDNA Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1693)
AUTHORS
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.

Location/Qualifiers

```

source.
1. .1693
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosomes="4"
/clones="RAFL15-01-B10 (R20394)"
/notes="This clone is in a modified pBluescript vector2
(lambda PS) as a BamHI/XhoI insert.
ecotype: Columbia"
1. .1693
/gene="At4g37320"
5'UTR
1. .27
/gene="At4g37320"
misc_difference 1. .21
/gene="At4g37320"
/notes="not present in genomic sequence"
28. .1515
/gene="At4g37320"
CDS
/codon_start=1
/evidence=experimental
/product="putative cytochrome p450 family protein"
/protein_id="AA022691.1"
/db_xref="GI:27754487"
/translation="MDLTQILLISFLPTISIKILLTKSNRKNLPSPNYPPLVPIGH
LHLKQPVHRTFHSISKSLGNAPIFHLRGLNRLVYVISHSIAEBEFTNDVVLNRP
DIIMAKHVGNFTNMAASYGDHNRNLRIRAAVEIFSHRISTFISKIDIRLLIITH
LSRSLHGFVEVELSKLLTNLAFNNIIMVAGKRYGTGTEDNDEAKLVRELIETHA
GAGSGLNADYLPVSNVNTFENOTKILGNLDRVLOKLVDEKRAEKEKGTLLIDHLS
FOETPEVYDVLIKGILALVLAGTDTSSVLEWMSNLNHPLEILKARABIDDKI
GSRDLVESIDVNLHYLQNVSETLRLYPVLPPLLPHFSSDECKVAGYDMPRTLILT
NVAMHRDPGLWEEPEKFEKGEARKLMPFGMRACPGABELKRLVSLALGC
LIOSFEWERYGAEVLVDTEGEGITMPKATPLRACAKAIVGKTI"
1516. .1693
/gene="At4g37320"
3'UTR
misc_difference 1678
/gene="At4g37320"
/notes="not present in genomic sequence"
BASE COUNT 482 a 394 c 374 g 443 t
ORIGIN
Query Match 21.2%; Score 394.8; DB 8; Length 1693;
Best Local Similarity 56.9%; Pred. No. 1.6e-78;
Matches 817; Conservative 0; Mismatches 592; Indels 27; Gaps 4;
Qy 162 CTCCTCTCTTCCCATATAGGAAACCTTAACTCTCGAACACGCCCAATCCACCGTTCT 221
Db 134 CTTACCCCTTTACCGGTTATCGGCACCTCCACCTCTCAAGCAACCGGTCCACCGACAT 193
Qy 222 TCCAACGGATCGAAGACAGTACGGCAACG-----TGGTTCCCTCTGGTTCCGTTAC 275
Db 194 TCCACTCCATCTCTAAGTCCCTTGGAAATGCTCCAATCTTCCACCTCCGCTGGGAAACC 253
Qy 276 GTCTGGCGGTTGTCTATCTCTCCACAGCATACCAAGATGCTTCACCAACACGACG 335
Db 254 GCCTCGTTTACGTATCTCTCACTCACTCATAGCGGAAGAATGTTTACCAAAACGACG 313
Qy 336 TTGCTTGGCCACCGGTACTCTCTCTCGGAAATAATACATCTTCTACAAACACCA 395
Db 314 TCGTTCTTGGAAACCGTCTTGATATCATCATGCGCCAAACACGTCGGCTATAACTTTACCA 373
Qy 396 CCGTAGGCTCTGTCCACGGCGAGCACTGGCGCACTCGCGGCATACCGCCCTGG 455
Db 374 ATATGATTTGCAAGCATCTTACGGCGACCACTGGAGGAATCTCCGCGCATCGCCCGCTG 433
Qy 456 ACGTCTCTCCAGCGAGCGGTCCACTCTCTCTCGGAATCCGGAGCGACGACGAAGC 515
Db 434 AGATATTTCTTCTCATAGATCAGTACCTTTCTTCTTATCCGTAAAGACGAGATCCGAC 493
Qy 516 GTCTGATCAGAGGTTGGTGTGGCCCAAGAACTCGAACGAGGAAGAGTTTCGCGAGTGG 575
Db 494 GGCTATAACACATCTCTTAGAGAC-----TCTTTCGACGAGTTTGTAGTGG 544
Qy 576 AGATTAGTTCGATGTTCAACGACTTAACCTTACAAACATTAATGAGGATGATATCGGGGA 635

```

RESULT 9
BT004449
LOCUS

Arabidopsis thaliana clone U20394 putative cytochrome p450 family
protein (At4g37320) mRNA, complete cds.

1519 bp mRNA linear PLN 14-FEB-2003

545 AGTTCGAAATCGTTATTAAACCACTTGGCAATCAACACATCATCATGATGGTAGCGGAA 604
Qy 636 AGAGGTTTTACGGAGAGGAGTAGATGAAGAACGTTGAGGAACCGGAGGAGTTTCAGAG 695
Db 605 AACGATATATACG-----TACCGGTACAGAGACAATGATGAAGCCAGCTCGTAGGG 658
Qy 696 AGACTGTGACAGAAATGTTGGAACCTCATGCGGTGCTGCTTAACAGGGAATCATCTTGCCTT 755
Db 659 AACTTTATAGCGAGATAATGCGCGCGCGGTTCTTGGAAATTTTGGCTGATTATCTTCCGT 718
Qy 756 TCCTCAGGTGGTTCCGATTTTTCAGATGTTGGAGAGCGCTTAAAGAGTATCAGTAAGAGGT 815
Db 719 CCATAAATTTGGGTC-----ACAAACTTTTGAGAACACAGACCAAAATCTTGGGAATCGAC 772
Qy 816 ACGATTTCATCTTGAATAGATCCTTTCATGAGAACCGTCCGACGCAATGACCCGCAAGATT 875
Db 773 TCGATAGAGTCTCTGCAAAAACCTGGTTGATGAGAAACGTGCAGAGAAAGAAAGGGTCAAA 832
Qy 876 CCATGATGATCATCTCTCCTCAAACTGCAAGAGACCCAGCCTCAGTACTACATGACCAAA 935
Db 833 CTTTGATCGATCATTTTGGCTTTCTTCCAGAAACCGAACCCGAGTACTATATCTGATGTC 892
Qy 936 TCATCAAAAGGCTTCTCTCTGCGCATGCTTTTGGTGGAACTGACTCATCACTGAGACTT 995
Db 893 TCATCAAAAGGAATCATATCTGGCTTTTGGTGGGAGACAGATACATCTTCACTGACGT 952
Qy 996 TAGAGTGGTCAATATCTAATTTATTTGAATCACCCAGAGAGTGTGAAGAGCAAGAGATG 1055
Db 953 TGGAAATGGCAATGTCANATTTGCTGACCATCCAGAGATATCTTGAAGAAAGCGAGCGG 1012
Qy 1056 AATTGGACACTCAAGTGGGCAAGACCGCTGTTTAAATAGTTCAGACTTCCAAAACCTTC 1115
Db 1013 AGATCGATGATAAATCGGTTTCAGACCCGTTAGTTGAAGAAATCAGATATTTGAAATCTCC 1072
Qy 1116 CATATCTTAGGAAGATCATCTTTCAGACACTTAAAGTTGTACCCCGCGGCCCAATTTCTAA 1175
Db 1073 ATTATCTCAAAACCATTTGTTCAGAAACATTAAGTTGTATCTCTGCGGTTCCACTACTAC 1132
Qy 1176 TACCTCATGTCTTTCAGAGATATTTACAATTTGAAGGATTCATATCCACGAGACACAA 1235
Db 1133 TCCCTCATTTCTCATCGGATGAATGTAAAGTGGCGGTACGATATGCCACGTCGCACGT 1192
Qy 1236 TTGTGATCATTAATGTTGGGCGATGACAGAGATCTCAGTTGTGGAATGATGCCACAT 1295
Db 1193 TGTATTATAAACAACGATATGGCGGATGATAGAGATCCAGGTTTATGGGAAGACCCAGAGA 1252
Qy 1296 GCTTTAAACCTGAGAGGTTTGTATGTGAAGAGAGAGAGAAAGTTGGTAGCATTTGGCA 1355
Db 1253 GGTTCAGCGGAGAGGTTTCGAGAAAGAGAGAGGCTCGAAAGCTAATGCCGTTTGGGA 1312
Qy 1356 TGGGAAGAGAGGCTTCCCGAGAGAACCCATGGCTATGCAAAAGTGTGACGTTTACCTTTGG 1415
Db 1313 TGGGACGACGAGCTTGTCTGGAGCTGAGCTTGGGAAGCGGTTAGTGAGCCTTGTCTCTTG 1372
Qy 1416 GATTGTTGATTCATGTTTGTGACTGGAACAGAGTAAGTAGGAGGAAGAGCTGTATATCAG 1475
Db 1373 GGTGCTGTGATTCAGTCTTTTCGAGTGGGAGAGAGTTGGTGCAGAACTTTGAGCATGACTG 1432
Qy 1476 AGAAACAATTTGGATCAGCTTGTCAAGGTTAATTCATTGGAGGCGCATGTGCAAGGCTCGCC 1535
Db 1433 AAGCGAAGGAGATCATATGCTTAAAGCTACTCCGTTGGAGCTATGTGCAAGGACGCTG 1492
Qy 1536 CACTTCCCACTAAAATTGGAATTTTAAATTAATAGTATTTTATTTTATTTTGGTAAACTT 1591
Db 1493 CCATTGTTGGTAAACGATATAATTTGACGCTATGTAGTACTATTTGTGTAGAAACCT 1548


```

Db      986 AGATCGATGATAAAATCGGTTTCAGACCGGTTAGTTGAAGAATCAGATATTTGTAATCTCC 1045
Qy      1116 CATATCTTAGAGATCATCTTGTAGACACTTAGTTGTACCCCGCCGCCCCCAATTTCTAA 1175
Db      1046 ATTAATCCAAACATTTGTGAGAACATTAAGTTTGTATCTCTGCGGTTCCACTACTAC 1105
Qy      1176 TACCTCATGTCTTCAGAAAGATTAATACAATGAAGGATTCATATATCCACGAGACACAA 1235
Db      1106 TCCCTCATTTCTCATCGATGAATGTAAAGTGGCGGCTACGATATGCGACGTCGACGTT 1165
Qy      1236 TTGTGATCATTAATGTTGGGCGATGAGAGAGATCTCTAGTTGTGGAATGATGCCACAT 1295
Db      1166 TGTATTAAACAAAGTATGGCGGATGATAGAGATCCAGGTTTATGGGAAGAGCCAGAGA 1225
Qy      1296 GCTTTAAACCTGAGAGGTTTGTATGGAAGGAGAGAGAAAGTTGGTAGCATTTGGCA 1355
Db      1226 GGTTCAGCGGAGAGGTTTCAGAAAGAGAGAGGCTCGAAAGCTTAATGCCGTTTGGGA 1285
Qy      1356 TGGGAAGAGGCTTGCACGAGAACCCATGCTATGCAAGTGTGAGCTTTACTTTGG 1415
Db      1286 TGGGACGACGAGCTTGTCTGGAGCTGAGCTTGGGAAGCGTTAGTGAGCCTTGCCTTG 1345
Qy      1416 GATTGTTGATTCATATGTTTGTGCTGGAACGAGTAAGTAGGAGAAAGCTTGATGACAG 1475
Db      1346 GGTGCTTGAATTCAGTCTTCAGTGGGAGAGAGTTGGTGCAGAACTTGTGACATGACTG 1405
Qy      1476 AGAACATTCGATCCTGCTCAAGGTTAATTAATTAATAGTATTTATTTATTTGGTAAAC 1589
Db      1406 AAGCGAAGGATCAGTATGCTTAAGCTACTCGGTTGCGAGCTATGTCGAGGCGAGTG 1465
Qy      1536 CACTTGGCCACTAAATTTGGAATTTAATTAATTAATAGTATTTATTTATTTGGTAAAC 1589
Db      1466 CCAATGTTGGTAACGATATAATTGACGTATGTAGGATCTATTGTGTAGAAAC 1519

RESULT 10
AX506493 1488 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 1188 from Patent WO0216655.
DEFINITION AX506493
ACCESSION AX506493
VERSION AX506493.1 GI:23387730
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1
REFERENCE
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1188 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source Location/Qualifiers
1..1488
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 408 a 360 c 336 g 384 t
ORIGIN
Query Match 21.2%; Score 393.8; DB 6; Length 1488;
Best Local Similarity 57.4%; Pred. No. 2.6e-78;
Matches 801; Conservative 0; Mismatches 567; Indels 27; Gaps 4;
Qy 162 CTCCTCTCTTCCCATATAGAAACCTTAACTCTCTCGAACACCCCAATCCACCGTTTCT 221
Db 107 CTTACCCCTTTACCGGTTATCGGCCACCTCCACCTCTCTCAAGCAACCCGCTCCACCGAAT 166
Qy 222 TCCAACGATGTCGAACAGTACGGCAACG-----TGGTTCCCTCTGTTGGTTCCAC 275

```

```

Db      167 TCCACTCCATCTCTAAGTCCCTTGGAAATGCTCCAATCTTCCACCTCCGCTTGGAAACC 226
Qy      276 GTCTGGCGGTTGTCTCATCTCTCTCTCAACAGACATACCAAGAATGCTTCCACCAACACGAGC 335
Db      227 GCCTGTTTACGTCATCTCTCTCACTCCATAGCCGAAGATGTTTCCACCAAAACGAGC 286
Qy      336 TTGCTTGGCCAAACCGGTACTCTCTCTCGGGAATAATACATCTTCTACAAACAACACA 395
Db      287 TCGTTCTTGGCAACCGTCTGATATCATATGCGCAACACATGCGCTGCTATTAATCTTTACCA 346
Qy      396 CGGTAGGCTCTGCTCCACCGCGAGCACTGGCGCAACCTCCGCGGCAATCACCGCCCTGG 455
Db      347 ATATGATGTCAGCATCTTACCGGACCACTGGAGAAATCTCCGCGCATTCGCGCGCTGC 406
Qy      456 AGCTCTCTCCACGCGAGCGGCTCACTCTCTCCGGAATCCGAGCGACAGAGCAAGC 515
Db      407 AGATATTTCTCTCATAGATCAGTACCTTTCTCTTATCCGTAAAGAGAGATCCGAC 466
Qy      516 GTCTGATGACAGAGTTGGTGGCCAAAGAACTCGAAAGAGAGAGTTTGGCCGAGTGG 575
Db      467 GGCTCATAAACATCTCTCTAGAGACT-----CCTTGACCGAATTTGTTGAAGTGG 517
Qy      576 AGATTAGTTCGATGTTCAACGACTTAACTTACAAACAATATAGGATGATATCGGCGA 635
Db      518 AGTTGAAATTCGTTATTAACCAACTTGGCAATCAACAACATCATCATGATGGTAGCCGAA 577
Qy      636 AGAGGTTTTACGAGAGGAGTGTGATGCAAGAAAGTTGAGGAAGCGAGGAGTTTCAGAG 695
Db      578 AACGATATTCAGG-----TACCGGTACAGAGACATGATGAGCCAGCTCGTGGGG 631
Qy      696 AGACTGTGACAGAAATGTTGGAACTCATGGGCTTGGCTTAAACAAGGAGATCATCTTGCCCT 755
Db      632 AACTTATAGCGAGATTAATGGCGCGCGGCTTCTGGAAATTTGGTGATTTATCTTCCGT 691
Qy      756 TCTCAGGTGGTTCGATTTTCAAGATGAGAGAGCGCTTAAAGAGTATCAGTTAGAGGT 815
Db      692 CATATAATTTGGGTC-----ACAAACTTTGAGAACAGACCAAAAATCTTTGGGGAATTCGAC 745
Qy      816 AGGATTTCATCTTGAATTAAGATCTTTCATGAGAACCGTGCAGCAATGACCCGCAAAAT 875
Db      746 TCGATAGAGTCTTCGCAAAAACCTGGTTGATGAGAAACGTCGACAGAAAGAAAGGGTCAAA 805
Qy      876 CCAATGATCATCATCTCTCAAACTGCAAGAGACCCAGCTCAGTACTACTACATGACCAAA 935
Db      806 CTTTGATCATCATTTGCTTTCTTCCAGAAACCGAACCCGAGTACTACTACTGTGCTCA 865
Qy      936 TCATCAAGGCTTCTCTGCGCATGCTTTTGGTGGACTGCTGCTCATCACTCACTGGGACTT 995
Db      866 TCATCAAGGAATCATACTGCGCTTTGGTCTTTTGGTTCTTTGCGGGGACAGATACATCTTCA 925
Qy      996 TAGAGTGGTCAATATCTAATTTTATTAATCAATCAATGAGAGTGTGGAAGAGGCAAGAGATG 1055
Db      926 TGGATGGCAATGTCAAAATTTGCTGAACCATCCAGAGATACCTTGAGAAAGCGAGAGCGG 985
Qy      1056 AATTGACACTCAAGTGGGCAAGACCGCTTGTAAATGAGTCAGACCTTCCAAACTTC 1115
Db      986 AGATCGATGATAAATTCGGTTTCAGACCCGTTAGTTGAAGAATCAGATATTTGTAATCTCC 1045
Qy      1116 CATATCTTAGGAAGATCATCTCTTGAGACACTTAGTTGTACCCCGCGGCCCAATTTCTAA 1175
Db      1046 ATTATCTCCAAAACATTTGTGTGAGAAACATTTAGTTTGTATCTCTGCGGTTCCACTACTAC 1105
Qy      1176 TACCTCATGTCTTCCAGAAAGATATTAACAATTAAGGATTTCAATATATCCACGAGACACAA 1235
Db      1106 TCCCTCATTTCTCATCGGATGAATGTAAAGTGGCGGCTACGATATGCCACGCTCGCACGT 1165
Qy      1236 TTGTGATCATTAATGTTGGGCGATGACAGAGATCTCTCAGTTGTGGAATGATGCCACAT 1295
Db      1166 TGTATTAAACAAACGATATGGGCGATGATAGAGATCCAGGTTTATGGGAAGAGCCAGAGA 1225
Qy      1296 GCTTTAAACCTGAGAGGTTTGTGTGGAAGGAGGAGAAAGTTGGTAGCATTTGGCA 1355
Db      1226 GGTTCAGCGGAGAGGTTTCGAGAAAGAGAGGCTCGAAAGCTTAATGCGGTTTGGGA 1285

```


Db 497 TAAATTTGCTAG-----CGAAAAAATCTCAACATGGGTTTGTGAAAGTGGAGATGA 547
Qy 582 GTTCGATGTTCAACGACTTAACATCAACACATAATGAGATGATATCGGGAAGAGGT 641
Db 548 AACCATTTGTTATGTTTAACTATCAACACATCATTTAGAAATGGTGGCGGAAGAAGAT 607
Qy 642 TTTACGGAGGAGAGTGAAGAACGTTGAGGAGCGAGGAGTTCAGAGAGACTG 701
Db 608 TCTACGGTGTGAACAG-----AGAACGACACAGAGGCTTAACACGTCGAGACAGCTAA 661
Qy 702 TGACAGAAATGTTGGAATCATGAGCTTGGCTTAACAAGGAGATCATCTTGCCTTTCTCA 761
Db 662 TTGCGGAGGTAGTTGTTAGCGCGCGCGAGGAAACGCGAGCAGACTCTCCCTATCCTTC 721
Qy 762 GGTGTTTCGATTTTCAGATGTGAGAGCGCTTAAAGAGTATCAGTAAGAGGTACCATTT 821
Db 722 GTT-----ACGTAACGAACTATGAGAAACATGTTAAGAGTTAGCAGGTGCTGTTGATG 775
Qy 822 CCATCTTTGAATPAAGATCCTTCATGAGAACCGTGCAGCAATGACCGCAAAATTCATGA 881
Db 776 AGTTTTTGCAAAGTCTAGTTAAACGAGAAACGTTGGAGAAAGTAAAGGTAACACAAATGA 835
Qy 882 TCGATCATCTCTCAACTCAAGAGACCGAGCTCTAGTACTACATGACCAAAATCATCA 941
Db 836 TCGATCATTTTACTTCTCTACAAGAAACTCAGCCCTGATTACTACACGGATGTGATCATCA 895
Qy 942 AAGGCTTGTCTCGCCCATGCTTTTGTGTGAATCTGCTCATCACTCACTCACTGGGACTTTAGAGT 1001
Db 896 AAGAAATCATACTTGTATTGATGTTTGTGCGGAGCTGATACATCACTGAGTGGAACTTAGAAT 955
Qy 1002 GGTCAATTAATTAATTTGAATCACCAGAGAGGTGTTGAAGAGCGCAAGAGATGAATTTGG 1061
Db 956 GGGCGATGCGAATTTGTTGAACCATCCAGAAATATTACGAAAGCGCAAGACCGAAATCG 1015
Qy 1062 ACATCAAGTGGGACAGACCGCTTGTAAATGAGTCAGACTTCCAAAATTCATATC 1121
Db 1016 ACGACCAATCGGTGTAGACCGGTGTAGAGAGGACAGACATTTGTGAACATCTCCTTACC 1075
Qy 1122 TTAGAAGATCATCTCTGAGACACTTAGTTGTACCCCGCCCGCCCAATTCATATACCTC 1181
Db 1076 TTCAGCATGTTGTCAGAGACTTTACGGCTTTATCCGGTGGCTCCAAATGCTTTTACCCC 1135
Qy 1182 ATGTGCTTCAGAGATATTCAATTTGAAGATTCATATATCCACGAGACACAAATGTGA 1241
Db 1136 ACTTAGCATCAGAAATTTGATAGTTGATGCTATGACGTTCCACGCGCACAAATCATCT 1195
Qy 1242 TCATTAAATGTTGGGCGATGAGAGATCTCTAGTTGTGGAATGATGCGACATGCTTTA 1301
Db 1196 TGGTGAATGCAATGGGCGCATCATAGATCCAAAGTTATGGGAAGAGCCAGAAAATTTA 1255
Qy 1302 AACCTGAGAGGTTTGTATGTGAAGAGAGAGAGAAAAGTTGGTAGCAATTTGGCATGGAA 1361
Db 1256 AGCCGAGAGGTTTGAGAAAAAAGAGAGAGGATAAAAGTTGATGCCATTTGGGATGGAC 1315
Qy 1362 GAAGGCTTGGCAGGAGAACCATGGCTATGCAAAAGTGTGAGCTTTACTTTGGGATGTT 1421
Db 1316 GACGATCTTGGCCTGGTTTCAAGGCTAGCTCAACGCTAGTGAATTTGGCTCTTGGATCAT 1375
Qy 1422 TGATTCATGTTTGTGACTGGAACAGTAAGTGAAGGAAAAGCTTGATATGACAGAGA--- 1478
Db 1376 TGGTCAATGTTTGTGAGTGGAGAGGGTTGAAGAAAAGTATCTAGACATGAGAGAAAGTG 1435
Qy 1479 ACAATTTGATCACTTGTCAAGGTTAAATTCATTTGGAGGCGCATGTGCAAGGCTCGCCAC 1538
Db 1436 AAAAAGGAACCAATGCGAAAAGCTACATCGTTGCAAGCTATGTGTAAGCTCGTCCCA 1495
Qy 1539 TTGCCACTAAATTTGAATTTAAATTAATTAATTAATTTTAT 1580
Db 1496 TTGTCCTAAGGTTTTAGATGCTTCTCTCTCCATACTTTTCAT 1537

RESULT 12
AY050849

LOCUS
DEFINITION

Arabidopsis thaliana putative cytochrome P450 protein (At3g28740)

AY050849 1612 bp mRNA linear PLN 18-SEP-2002

complete cds.

AY050849

AY050849.1 GI:15292830

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 1612)

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,

Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,

Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,

Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,

Davis,R.W., Ecker,J.R. and Theologis,A.

Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 1612)

Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M.,

Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L.,

Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,

Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M.,

Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,

Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and

Theologis,A.

Direct Submission

Submitted (08-AUG-2001) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN

Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,

Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,

Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the

sequencing and annotation of the RAPL cDNAs: Yamada,K., Liu,S.X.,

Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D.,

Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,

Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H.,

Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B.,

Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,

Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGECC) and Seki,M. (RIKEN GSC) contributed equally to

this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGECC)

contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis

genome submitted to GenBank.

Location/Qualifiers

1. .1612

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAPL08-11-J17 (RI1222)"

/note="This clone is in a modified pBluescript vector

(FLC-1) as a BamHI/XhoI insert.

ecotype: Columbia"

1. .1612

/gene="At3g28740"

1

/genes="At3g28740"

2. .1531

/gene="At3g28740"

/codon_start=1

/evidence=experimental

FEATURES
source

gene

5'UTR

CDS

JOURNAL Patent: WO 0216055-A 2278 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)

FEATURES
source Location/Qualifiers
1..1494
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 399 a 341 c 355 g 399 t
ORIGIN

Query Match 20.8%; Score 386.6; DB 6; Length 1494;
Best Local Similarity 57.0%; Pred. No. 1.1e-76;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

QY 90 TTTTCTCGGTCTATTCTCGCGGTGAAGTTTGTTCCTTCCAAAGCAGAAAAATTGAGAAACA 149
DB 29 TTCTCTCTGTTGTTCTCTCTCTCAATTACTTAAATGGAAACTCAAGCGAAAGCCAAATC 88
QY 150 TACACACAGGTCCTCC---TCCTCTTCCCATAATAGGAAACCTTAACTCTCTCGAACAGC 206
DB 89 TACCTCCGAGTCGGCATATGCGGTGATTGGTCATCTCCGCTTCTCAAAACCAAC 148
QY 207 CAATCCACCGTTTCTTCCACGCAATGCGAAACAGTAGCGCAACG-----TGGTTTCCC 260
DB 149 CGATTTCATCGCACATCTCTCCCTCTCTCAATCCCTTAAACAAATGCTTCGATCTTCTCCC 208
QY 261 TCTGTTTGGTTTACGCTGCGCGTGTGTCATCTCTCTCCAAAGCATATCAAGAAATGCT 320
DB 209 TCCGACTCGGTAACCGACTGTTTTGTAACCTCGTCACTCGATCGCGAGGATGTT 268
QY 321 TCACCAACACGAGCTTGCCTTGGCCAAACCGGCTACCTTCTCTCGGGAATAATACATCT 380
DB 269 TCACCAAAACGAGCTGTACTGCGGAACAGACCAAACTTCTCTCGTAAACACGCTG 328
QY 381 TCTACAAACACACCGTAGGCTCTGCTCCACGCGGAGCACTGCGGCAACCTCGCC 440
DB 329 CGTAGCATACACAAACCATGATCGAGCTTCTTCAAGTGACCACTGCGGTAAACCTCGCC 388
QY 441 GCATCACCGCTCGAGCTCTCTCCAGCAGCGCGTCCACTCTCTCGGAAATCCGGA 500
DB 389 GCATCGGCTCGTCAGATATCTCCATCACCGTCTCAATAGCTTTTGTCTATTCGTA 448
QY 501 CGCAGCAGACGAAGCGTCTGATCAGAGGTTGGTGGCCAGAACTCGAACGAGGAAG 560
DB 449 AAGACGAGATCCGACGACTTGTGTTCTCTTCTTCGGAAC-----TTTTTCAACAG 499
QY 561 AGTTTGGCGAGTGGAGATTAGTTCGATGTTCAACGACTTAACTTACAAACAAATATGA 620
DB 500 AGTTTGGAAAGTGATGAATGAATCAATGTTATCTGACTTAACTTCAACAACTTTTAA 559
QY 621 GGATGATATCGGGGAAGAGGTTTACGGAGAGGAGGTGAGATGAAGAACGTTGAGGAAG 680
DB 560 GAATGGTGGCGGAAACGTTTACTACGGAGAGCGGTGTTG-----AGGATGATCCGGAGG 613
QY 681 CGAGGAGTTCAGAGAGACTGTGACAGAAATGTTGGAACTCATGGCTTGGCTTAAACAGG 740
DB 614 CTAACCGTGTCCGCGAGCTTATAGCGGATGTGGTGGCTTGTGCTGTGCTGGAACGCTG 673
QY 741 GAGATCACTTCCCTTCTCAGGTGTTTCGATTTTCAGAAATGTCGAGAGCGCTTAAAGA 800
DB 674 TTGATTACTTACCGGTTTTCGGTGG-----TTTCAGATTACGAGACCGGGTTAAGA 727
QY 801 GTATCAGTAAAGAGTACGATTCCATCTTGAATAAGATCTTTCATGAGAACCGTCCAGCA 860
DB 728 AGTTGGCGGTAGGCTCGACGAGTCTTTCGAGGATTGGTTGATGAGAAACGAGAAGCTA 787
QY 861 ATGACCGCCAGAAATTCATGATGATGATCTCTCAAACTCGAAGACCCAGCCCTCAGT 920
DB 788 AGGGAAGAGGAAACACTATGATCGATCACTTGTCTTCTGCAAGAAATCCCAACCGGATT 847
QY 921 ACTACACTGACCAATCATCAAGCGCTTGTCTGGCCATGCTTTTGGTGGACTGACT 980

DB 848 ACTTACCGATCGTATCATTTAAAGAAACATGCTCGCTTGTATCTAGCAGGACCGACA 907
QY 981 CATCACTGGGACITTTAGAGTGGTCAITATCTAATTTATTGAATCAACCCAGAGGTGTGA 1040
DB 908 CATCAGCGGTACGTTAGATGGGCAITGTGCAACGCTGTGAACCATCCGATGTATTGA 967
QY 1041 AGAAGCAGAGATGAATGGACATCAAGTGGGACAGACCGCTTGTTAATGATGATCAG 1100
DB 968 ACAGGCGAGAGATGAATCGATGAAGATAGGTTTACACAGGCTTATGGATGAATCAG 1027
QY 1101 ACCTTCCAAACTTCCATATCTTAGAAGATCATCTTGGAGACACTTAGTGTGTACCCCC 1160
DB 1028 ATATCTCAAACTGCTTATCTCCAAACATTTGTCTGAAACGTTGCGCTTTATCTCTG 1087
QY 1161 CGGCCCCAATTCAATACCTCAATGCTTTCAGAAAGATTAACAATTAAGGATTCATA 1220
DB 1088 CGGCTCCCATGCTTCTCTCCTCAGTTGCTCGAAAGATTGTAAGTTGCAGGATACGATA 1147
QY 1221 TCCACGAGACACAATTTGATCATTAATGTTGGGGCATGCGAGAGATCTCTCAGTTCT 1280
DB 1148 TCCCGGTGGCAGATCTATTGACCAATGTTGGGCTTATACAGAGATCTCTCAGCTAT 1207
QY 1281 GGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTGTAAGAGAGAGAGAAAGT 1340
DB 1208 GGGATGATCCATGAGCTTCAAGCCAGAGAGGTTTGAGAAAGAGAGAGCTCAGAGC 1267
QY 1341 TGGTAGCAATTTGGCATGGGAAGAGGCTTGCAGGAGAACCCATGGCTATGCAGAAAGT 1400
DB 1268 TAATGCGGTTTGGTTAGGAAGAGGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1327
QY 1401 TCAGCTTTACTTTGGGATTTGATTCATGTTTGTGCTGGAACGAGTAAGTGAGGAA 1460
DB 1328 TAAACCTGACTCTTGGATCAATGATTCAGTGTTCGAAATGGGA---GAAGATTTGGAGAAG 1384
QY 1461 AGCTTGATATGACAGAGAACTTGAATGATCCTTGTCAAGGTTAAATTCATTTGGAGGCCA 1520
DB 1385 AGTGATATGATGAGCAAGAGTGTTCATGCTTAAAGCCAAAGCCCTTTGGAGGCCA 1444
QY 1521 TGTGCAAGCTCGCCACATTTGCCACTAAAT 1551
DB 1445 TGTGCAAGCAGCTCCCTCTCTGTTTAAAT 1475

RESULT 14
AX590025
LOCUS 1494 bp DNA linear PAT 24-JAN-2003
DEFINITION: Sequence 207 from Patent WO02081695.
ACCESSION: AX590025
VERSION: AX590025.1 GI:27901173
KEYWORDS: Arabidopsis thaliana (thale cress)
SOURCE: Arabidopsis thaliana
ORGANISM: Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS: Zhu, T., Glazov, E.A., Meins, F., Wang, X. and Chang, H.S.
TITLE: Genes that are modulated by posttranscriptional gene silencing
JOURNAL: Patent: WO 02081695-A 207 17-OCT-2002;
Syngenta Participations AG (CH)
FEATURES
source Location/Qualifiers
1..1494
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 399 a 341 c 355 g 399 t
ORIGIN

Query Match 20.8%; Score 386.6; DB 6; Length 1494;
Best Local Similarity 57.0%; Pred. No. 1.1e-76;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

QY 90 TTTTCTCGGTCTATTCTCGCGGTGAAGTTTGTTCCTTCCAAAGCAGAAAAATTGAGAAACA 149

Db 29 TTCTCTTGGTGTCTCTCTCAATTAAATGGAATACTCAAGCAAGAAACCAATC 88
QY 150 TACACACAGTCTCTCC---TCTCTTCCCATATAGGAACCTTAACCTCTCTCGAACAGC 206
Db 89 TACCTCGAGTCGGCATGCTGATACCGGTGATGCTCATCTCCGCTTCTCAACACAC 148
QY 207 CAATCCACCGTTCTTCCCAACGATGTGCAAAACAGTACGGCAACG-----TGGTTTCCC 260
Db 149 CGATTTCATGCGCATTTCTCTCCCTCTCTCAATCCCTAAACCAATGCTCGGATCTTCTCCC 208
QY 261 TCTGTTTCTGTTTCACTCTCGGCGTGTGCTCATCTCTCTCCACAGCATACCAAGATGCT 320
Db 209 TCGGATCGGTAACCGACTGTTTTCGTAACCTGTCACATCGATCGCGGAGATGTT 268
QY 321 TCACCAAAACAGACGTTGCTTGGCCAAACCGGCTACCTTCTCTCGGGAATAATCATCT 380
Db 269 TCACCAAAACAGACGTCGTAACCTGCGCAACAGACCAACTTCATCTCGCTTAACAGTTG 328
QY 381 TCTAACCAACACACCGTAGGTCCTGTCCTGTCCTCCACGGGAGACATCGGCAACCTCGGC 440
Db 329 CGTACGATTACACACCATGATCGCAGCTTCTCTACGGTGACCACTGGCGTAACCTCGGC 388
QY 441 GCATCACCGCTCGGAGCTCTCTCCACGAGCGCTGCCATCTCTCGGAACTCGGA 500
Db 389 GCATCGCTCGGTCGAGATATCTCCCAATCACCCTCTCAATAGCTTTTGTCTATTTCGTA 448
QY 501 GCGACGAGACGAACGCTGTGATCAGAGGTTGGTGTGCGCAAGAACTCGAACGAGGAAG 560
Db 449 AAGACGAGATCCGACGACTGTGTTTCGTTCTTCGGAAC-----TTTTCACAG 499
QY 561 AGTTTCGGGAGTGGAGATGATGTCGATGTTCAACGACTTAACATTAACAACATTAAGA 620
Db 500 AGTTTGTGAAGTGGATGAAATCAATGTTATCTGACTTAACATTAACAACATTTTAA 559
QY 621 GGATGATATCGGGAAGAGTTTACGAGAGAGATGAGATGAAGACGTTGAGGAG 680
Db 560 GAATGTGCGCGGAACAGTTTACTACGAGACCGTGTG-----AGGATGATCCGGAGG 613
QY 681 CGAGGAGTTTCAGAGACTGTGACAGAAATGTTGGAACCTCATGGCTTGGCTTAAACAAG 740
Db 614 CTAAACGTCGCGCAGCTTATAGCGATGTGTCGCTGTGCTGTAAGACGCTG 673
QY 741 GAGATCACTGCTTCTTCTCAGTGTGTCGATTTTTCAGAAATGAGAGACGCTTAAAGA 800
Db 674 TTGATTACTTACCGGTTTTCGCGTTGG-----TTTCAGATTACGAGACACGCGTTAAGA 727
QY 801 GTATCAGTAAGAGTACGATTCCTCTTGAATAGATCTTCAATGAGAACCGTGCAGCA 860
Db 728 AGTTGCGGTAGGCTCGAGAGTTCTTGAAGGATGTTGATGAGAAACGAGAAAGTGA 787
QY 861 ATGACCGCCAGAAATTCATGATGATCATCTCTCAAACTGCAAGAGACCCAGCTCAGT 920
Db 788 AGGAGAAGAAACACTATGATGATCATCTGTTTACTCTGCAAGATCCCAACCGATT 847
QY 921 ACTACACTGACCAATATCAAAAGCGCTGCTCTGGCAGATGCTTTTGTGGAACATGAT 980
Db 848 ACTTACGATCGTATCAATTAAGGAAACATGCTCGCTTTGATCTAGCAGGACCGACA 907
QY 981 CATCACTGGGACTTTAGAGTGTCTATCTAATTTATGATCAACCCAGAGGTTGCTCA 1040
Db 908 CATCAGCGGTTAGTATGAGATGGCAATGTCGAACCGTTGTAACCAATCCGATGATTA 967
QY 1041 AGAAGGCAAGAGATGAATGGAACATCAAGTGGGCAAGACCGCTTGTAAATGATGATCAG 1100
Db 968 ACAAGGAGAGATGAATGATAGAGAAAGATAGTTTGAAGAGCTTATGATGATCAG 1027
QY 1101 ACCTTCCAAACCTTCATATCTTGAAGAGATCATCTTGAAGACATTAAGGTTGACCCCC 1160
Db 1028 ATATCTCAAAACCTGCTTATCTCCAAACATTTGTGCTGAACGTTTTCGCTTTATCTCG 1087
QY 1161 CGGCCCCCAATTTCAATACCTCATGTTCTTCAGAAATATTAATTAAGGATTCATTA 1220

Db 1088 CGGTCCCATGCTTCTCTCTCACTGCTCGGAAGATTTGTAAGTTCGAGATACGATA 1147
QY 1221 TCCACGAGACACAAATTTGTGATCATTTAATGTTGGGGCATGACAGAGATCTCTCAATGCT 1280
Db 1148 TCCCGGTGGCAACGATATACTATTGACCAATGTTGGCTTATACACAGAGATCTCTCACTAT 1207
QY 1281 GGAATGATGCCACATGCTTTTAAACCTGAGAGGTTTGTGTTGGAAGAGAGAGAAAGT 1340
Db 1208 GGGATGATCCAAATGAGCTTCAAGCCAGAGAGGTTTGAAGAAAGAGAGAGAGCTCAGAAGC 1267
QY 1341 TGGTAGCATTTGGCATGGAAGAGGGCTTCCCAAGGAGAACCCATGCTATGCAAGATG 1400
Db 1268 TAATGCCGTTTGGGTTAGGAAGAGGGCGTGTCTGGTCTGGACTGGCTCATCGGCTTA 1327
QY 1401 TCAGCTTTTACTTTGGGATTTGATTCATGTTTGACTGGAAACGAGTAAGTGAAGAAA 1460
Db 1328 TAAACCTGACTCTTGGATCATTTGATTCAGTTGTTGAATGGA---GAAGATTGGAGAAG 1384
QY 1461 AGCTTGATATGACAGAGAAACAATTGGATCACCTTGTCAAGGTTAATTCATTTGGAGGCCA 1520
Db 1385 AAGTGATATGATGAAGGCAAGGTTTACAATGCTTAAAGCCAAAGCCTTTTGAAGCCA 1444
QY 1521 TGTGCAAGGCTCGCCACTTGCCTCACTAAAT 1551
Db 1445 TGTGCAAGGACAGTCCCTCTCTGTTGTTAAAT 1475

RESULT 15
AX651453 1494 bp DNA linear PAT 22-MAR-2003
LOCUS AX651453
DEFINITION Sequence 249 from Patent WO03000898.
ACCESSION AX651453
VERSION AX651453.1 GI:29154271
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Karagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 249 03-JAN-2003;
Syngenta Participations AG (CH)
Location/Qualifiers
1. 1494
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 399 a 341 c 355 g 399 t
ORIGIN

Query Match 20.8%; Score 386.6; DB 6; Length 1494;
Best Local Similarity 57.0%; Fred. No. 1.1e-76;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

QY 90 TTTTCTCTGTTCTATTCTCTCGGCGTGAAGTTGTTTCCAAAGCAGAGAAATTTGAGAAACA 149
Db 29 TTTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 88
QY 150 TACCAACAGGTCCTCC---TCTCTTCCCAATAATAGAAACCTTAACCTCTCTCGAACAGC 206
Db 89 TACCTCCGAGTCGGCATGCTTATTCGCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 148
QY 207 CAATCCACCGTTTCTTCCAAACGATGTGCAAGAGCTTAAAGAACTTAACCTCTCTCGAACAGC 206
Db 149 CGATTCATCGCAACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 208
QY 261 TCTGTTTCTGTTTCACTCTGCGCGTGTGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 320
Db 209 TCCGACTCGGTAAACGACTCTGTTTTCGTTGTAATCTGCTACATCTGATCGCGAGGATGTT 268

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 02:38:44 ; Search time 3601 Seconds
(without alignments)
12547.076 Million cell updates/sec

Title: US-09-868-546A-1

Perfect score: 1859
Sequence: 1 gaaacactgacagacagca.....aataaattttttactttc 1859

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	596.4	32.1	598	13	BQ786681 saq71h06.
2	582.6	31.3	589	12	BI321303 saf50c06.
3	573	30.8	573	13	BU579309 saf56f09.
4	491.8	26.5	600	13	BQ742415 saq43a09.

C	5	477.2	25.7	743	12	BI969319
	6	470.6	25.3	582	9	AW234443
	7	468	25.2	581	9	AW307234
	8	465.2	25.0	574	13	BU764794
	9	455.8	24.5	574	12	BI498223
	10	455.4	24.5	561	12	BM187538
	11	451.2	24.3	544	13	BQ079433
	12	444.6	23.9	564	13	BQ453555
	13	426	22.9	564	13	BU762419
	14	417	22.4	425	12	BI424885
	15	410.8	22.1	426	12	BI426262
	16	410	22.1	552	9	AW782116
	17	406.4	21.9	420	13	BQ742929
C	18	405.8	21.8	651	9	AW351157
	19	400.2	21.5	503	9	AI495626
	20	398.2	21.4	524	12	BI971995
	21	396.6	21.3	696	14	CB891448
	22	395.2	21.3	473	10	BG046583
C	23	383.6	20.6	666	9	AW309826
C	24	374.6	20.2	789	9	AW102198
	25	373.8	20.1	755	14	CB893434
	26	367.8	19.8	846	14	CB894274
	27	362.6	19.5	845	14	CB8931246
C	28	355.6	19.1	608	9	AW309498
	29	346.2	18.6	610	13	BQ612582
	30	342	18.4	422	13	BQ080894
	31	337.4	18.1	491	13	BU762460
C	32	329	17.7	787	10	BG644813
	33	328.8	17.7	784	14	CB893907
C	34	328.2	17.7	673	10	BG045616
	35	321.8	17.3	754	12	BM779857
	36	318.6	17.1	770	10	BG648962
	37	316.4	17.0	769	14	CB893509
	38	314.4	16.9	576	13	BU764619
	39	313	16.8	593	10	BE205034
	40	308.8	16.6	596	9	AW185361
	41	305.4	16.4	587	10	BE190273
C	42	301.8	16.2	797	14	CA922487
	43	297	16.0	712	9	AW775904
	44	290.2	15.6	723	14	CB893510
C	45	288.2	15.5	591	13	BU545044

ALIGNMENTS

RESULT 1

BQ786681

LOCUS

DEFINITION

saq71h06.y1 Cm-cl076

Glycine max cDNA clone

SOYBEAN CLONE ID:

Gm-cl076-5363 5' similar to TR:Q9ZWF2

Q9ZWF2 CYTOCHROME P450. 1,

mrna sequence.

ACCESSION

BQ786681

VERSION

BQ786681.1

KEYWORDS

EST.

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

REFERENCE

1 (bases 1 to 598)

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished

Public Soybean EST Project

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 421.

Location/Qualifiers

1. .598

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1076-5363"

/tissue_type="wounded cotyledons"

/dev_stage="11 day old seedlings"

/lab_host="DH10B"

/clone_lib="Gm-c1076"

/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 11 day old seedlings treated with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Iliia Vodkin lab at the University of Illinois)." 139 t

BASE COUNT

ORIGIN 143 a 180 c 136 g 139 t

Query Match

Best Local Similarity 32.1%; Score 596.4; DB 13; Length 598;

Matches 597; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAAACACTGACAGACGATAGTCTCTGGTGCAAGAAATCAATTGACGACATGGGAAT 60

Db 1 GAAACACTGACAGACGATAGTCTCTGGTGCAAGAAATCAATTGACGACATGGGAAT 60

Qy 61 GTTCTGGTGTGCTCTCATACGCTGCTTTTCTGTTCTATCTCTCGGCGTGAAGTT 120

Db 61 GTTCTGGTGTGCTCTCATACGCTGCTTTTCTGTTCTATCTCTCGGCGTGAAGTT 120

Qy 121 TGTCTTCCAAAGCAGAAATTGAGAAACATACACGAGGTCTCTCTCTCTTCCCATAT 180

Db 121 TGTCTTCCAAAGCAGAAATTGAGAAACATACACGAGGTCTCTCTCTTCCCATAT 180

Qy 181 AGGAAACCTTAACTCTCTGACAGCAATCCACCGTTTCTTCAACGCGATGCGAAACA 240

Db 181 AGGAAACCTTAACTCTCTGACAGCAATCCACCGTTTCTTCAACGCGATGCGAAACA 240

Qy 241 GTACGGCAACGTGTTTCCCTCTGGTTCGGTTCAGCTTCGCGGTGTGATCTCTCTCC 300

Db 241 GTACGGCAACGTGTTTCCCTCTGGTTCGGTTCAGCTTCGCGGTGTGATCTCTCTCC 300

Qy 301 AACAGCATACCAAGAAGTCTTCAACAAACAGAGGTTCGCTTGGCCCAACCGGTACCTTC 360

Db 301 AACAGCATACCAAGAAGTCTTCAACAAACAGAGGTTCGCTTGGCCCAACCGGTACCTTC 360

Qy 361 TCTCTCGGAAATACATCTTCTACAAACACACCGGTAGGCTCTGCTCCACGCGCA 420

Db 361 TCTCTCGGAAATACATCTTCTACAAACACACCGGTAGGCTCTGCTCCACGCGCA 420

Qy 421 GCACCTGGCGCAACCTCCGCCGATACCGCCCTCGACGTCCTCTCCACGAGCGGTCCA 480

Db 421 GCACCTGGCGCAACCTCCGCCGATACCGCCCTCGACGTCCTCTCCACGAGCGGTCCA 480

Qy 481 CTCCTTCTCCGAATCCGGAGCGACGAGACGAAGCTCTGATGCGAGAGTTGGTGTGGC 540

Db 481 CTCCTTCTCCGAATCCGGAGCGACGAGACGAAGCTCTGATGCGAGAGTTGGTGTGGC 540

Qy 541 CAAGAACTCGAAGAGAGAGTTTCCGCGAGTGGAGATTAGTTCGATTTCAACGAC 598

Db 541 CAAGAACTCGAAGAGAGAGTTTCCGCGAGTGGAGATTAGTTCGATTTCAACGAC 598

RESULT 2

BI321303

LOCUS

DEFINITION

BI321303.1

GI:15000489

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 599)

REFERENCE

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V., Khanna

, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished

CONTACT: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 421.

Location/Qualifiers

1. .589

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1077-2027"

/tissue_type="18 day old 'Williams' seedlings"

/dev_stage="18 day old 'Williams' seedlings"

/lab_host="DH10B"

/clone_lib="Gm-c1077"

/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The mRNA was isolated from cotyledons of 18-day-old

'Williams' seedlings which were greenhouse grown in

potting soil. The cotyledons were flash-frozen in liquid

nitrogen. Stratagene's cDNA Synthesis Kit (catalog number

200401) was used to synthesize the cDNA. First-strand

synthesis was performed with 5-methyl dCTP, hence the

ligated cDNA was hemimethylated. A modification of

Stratagene's first-strand synthesis primer was used. An

'anchor' nucleotide (V=A, C, or G) was added to the 3' end

of the primer (GAGAGAGAGAGAGAGAGTCTCGAG(T)18V) to

anchor the primer at the 5' end of the poly(A) tract.

After second-strand synthesis, the cDNA ends were filled

in with cloned Pfu DNA, ligated to EcoRI adapters and

subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The xhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all xhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl 8-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constructed in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University."

BASE COUNT 173 a 127 c 134 g 155 t
ORIGIN

Query Match 31.3%; Score 582.6; DB 12; Length 589;
Best Local Similarity 99.3%; Pred. No. 2.6e-120;
Matches 585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 748 CTTCGCTTTCCTCAGCTGGTTCGATTTTCAGATGTGGAGAGCCGCTTAAAGAGTATCAG 807
DB 1 CTTCGCTTTCCTCAGCTGGTTCGATTTTCAGATGTGGAGAGCCGCTTAAAGAGTATCAG 60

QY 808 TAAGAGGTACGATTCATCTTGAATAAGATCCTTTCATGAGAACCGTGCCAGCAATGACCG 867
DB 61 TAAGAGGTACGATTCATCTTGAATAAGATCCTTTCATGAGAACCGTGCCAGCAATGACCG 120

QY 868 CAGAAATTCATGATCGATCATCTCTCAAATGCAAGAGAGCCAGCCCTCAGTACTACAC 927
DB 121 CAGAAATTCATGATCGATCATCTCTCAAATGCAAGAGAGCCAGCCCTCAGTACTACAC 180

QY 928 TGACCAAAATCATCAAGAGCCCTGCTCTGCGCATGCTTTTGTGGAACTGACTCATCAAC 987
DB 181 TGACCAAAATCATCAAGAGCCCTGCTCTGCGCATGCTTTTGTGGAACTGACTCATCAAC 240

QY 988 TGGGACTTTAGAGTGCTATTATCTAATTTATGTAATCACCAGAGGTGTTGAAGAAGGC 1047
DB 241 TGGGACTTTAGAGTGCTATTATCTAATTTATGTAATCACCAGAGGTGTTGAAGAAGGC 300

QY 1048 AAGAGATGAATTGGACATCAAGTGGGAGCAAGACCGCTGTTAAATGATCAGACTTCC 1107
DB 301 AAGAGATGAATTGGACATCAAGTGGGAGCAAGACCGCTGTTAAATGATCAGACTTCC 360

QY 1108 AAAAATTCATATCTTAGAAGATCATCTCTGAGACATTAGGTGTACCCCGGCCCC 1167
DB 361 AAAAATTCATATCTTAGAAGATCATCTCTGAGACATTAGGTGTACCCCGGCCCC 420

QY 1168 AATTCCTAATACCTCATGTGCTCTTCAAGATATTCAATTAAGAGATTCAATATCCACG 1227
DB 421 AATTCCTAATACCTCATGTGCTCTTCAAGATATTCAATTAAGAGATTCAATATCCACG 480

QY 1228 AGACAAATTTGATCATTAATGTTGGGCGATGACAGAGATCTCTCAGTTGGAATGA 1287
DB 481 AGACAAATTTGATCATTAATGTTGGGCGATGACAGAGATCTCTCAGTTGGAATGA 540

QY 1288 TGCCACATGCTTTAAACCTGAGAGGTTTGTATGTTGAGAGAGGAGAAA 1336
DB 541 TGCCACATGCTTTAAACCTGAGAGGTTTGTATGTTGAGAGAGGAGAAA 589

RESULT 3
BUS79309
LOCUS
DEFINITION
sars6f09.y1 Gm-cl074 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl074-6425 5' similar to TR:Q9ZRW6 Q9ZRW6 CYTOCHROME P450 ;,
mRNA sequence.
ACCESSION BUS79309
VERSION BUS79309.1 GI:23064536
KEYWORDS EST.

SOURCE ORGANISM

Glycine max (soybean)
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE

1 (bases 1 to 573)

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished

COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: cu@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 425.

Location/Qualifiers

FEATURES

source

1..573

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clones="SOYBEAN CLONE ID: Gm-cl074-6425"

/tissue_type="seedlings induced for HR (hypersensitive response)"

/dev_stage="9-11 day old"

/lab_host="DH10B"

/clone_lib="Gm-cl074"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 173 a 99 c 135 g 166 t

ORIGIN

Query Match 30.8%; Score 573; DB 13; Length 573;

Best Local Similarity 100.0%; Pred. No. 3.7e-118;

Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 TGAATTGGACACTCAAGTGGGAGCAAGCCGCTTGTAAATGAGTCAGACCTTCCAAAAT 1113

DB 1 TGAATTGGACACTCAAGTGGGAGCAAGCCGCTTGTAAATGAGTCAGACCTTCCAAAAT 60

QY 1114 TCCATATCTTAGGAAGATCATCTTGAGACACTTAGGTGTGACCCCGGCCCAATTCT 1173

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

TITLE

Public Soybean EST Project

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert length: 959 Std Error: 0.00
High quality sequence stop: 423.

FEATURES

source

1. 582
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl028-989"
/tissue_type="roots of 'Supernod' plants"
/lab_host="DH10B"
/clone_lib="Gm-cl028"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of Glycine max 'Supernod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. Stratagene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A,C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAACTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT

ORIGIN

Query Match 25.3%; Score 470.6; DB 9; Length 582;
Best Local Similarity 88.7%; Pred. No. 4.2e-95;
Matches 509; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 859 CAATGACCGCAGAAATCCATGATCGATCTCTCAAACTGCAAGACCCAGCCTCA 918

Db 9 CAAGACCGCAGAAATCCATGATCGATCTCTCAAACTGCAAGACACACGCTGA 68

Qy 919 GTACTACACTGACCAAAATCATCAAGCGCTTGCTCTGCGCATGCTTTTGGTGAACGA 978

Db 69 CTATTATACCGACCAATCATCAAGACATTCCTTTGGCTATGCTCTTTGGCGAACA 128

Qy 979 CTCATCAACTGGAGACTTTAGAGTGTCTAATATCTAATTTAATGAATCAACCCAGAGGTGT 1038
Db 129 CTCGTCAACTGGAACTTTAGAGTGGCATTAATTAATTTAGTGAATGACCCAGAGGTGT 188
Qy 1039 GNAAGAGCGCAGAGATGAATTTGGACACTCAAGTGGCAAGACCGCTTGTAAATGAGTC 1098
Db 189 GCAGAAAGCAAGATGATGAGTTGGACGCTCAAGTAGGACAGATCGGCTGTTAAATGAGTC 248
Qy 1099 AGACCTTCCAAAACCTTCCATATCTTTAGGAAGATCATCTTTGAGACACTTTAGTGTGTACCC 1158
Db 249 AGACCTTCCAAAACCTTCTTATCTCAGGAAGATAGTCTTCTGAAACACTTTAGTGTGTACCC 308
Qy 1159 CCGGCCCCCAATTTAATACCTCATGTCTTTCAGAAAGATATTACAATGAAGATTCAA 1218
Db 309 TCCGGCTCCAATTTAATACCAACACGTGGCTTCAGAAGACATCAATATCGAAGGATTCAA 368
Qy 1219 TATCCACGAGACACAATTTGTGATCATTAATGTTGGGCGATGCGAGAGATCTTCAGTT 1278
Db 369 TGTTCACGAGACACAATTTGTGATTTAATGTTGGGCGATGCGAAGAGATCTTAAGAT 428
Qy 1279 GTGGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTGTTGGAAGGAGAGAGAAAA 1338
Db 429 ATGGAAGATGCGACAAGCTNTAAACCTGAGAGGTTTGTGTTGGAAGGAGAGAGAGAA 488
Qy 1339 GTTGTGATCATTTGGCATGGGAAGAGGCTTCCCGAGAGAACCCATGCGCTATGCAAG 1398
Db 489 GTTGTGATCATTTGGTATGGAAGAGGCTTCCCGAGAGAACCCATGCGCTATGCAAG 548
Qy 1399 TCTGAGCTTACTTTGGGATTTGCTGATTCAATGT 1432
Db 549 TCTTGTATATCTTTGGGATTAATGATTCAATGT 582

RESULT 7

AW307234

LOCUS

DEFINITION

sf54d12.y1 Gm-cl009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: 'Gm-cl009-3792 5', similar to TR:Q9ZRW6 Q9ZRW6 CYTOCHROME P450 ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AW307234 581 bp mRNA linear EST 02-DEC-2001
AW307234.1 GI:6719587
EST
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 581)

REFERENCE

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

Unpublished
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert length: 924 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 422.

FEATURES

source

1. 581
/organism="Glycine max"

Query Match 24.5%; Score 455.8; DB 12; Length 574;
 Best Local Similarity 86.9%; Pred. No. 8.9e-92;
 Matches 499; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 826 CTTGATAGATCCCTCATGAGAACCGTCCGACGATGACCCGAGAAATTCATCATCGA 885
 DB 1 CTTGATAGATCCCTCATGAGAACCGTAAACACAGGACCGGAGAAATTCATCATCG 60

QY 886 TCATCTCTCCCAAACTGCAAGAGACCCAGCGCTCAGTACTACACTGACCAATCATCAAGG 945
 DB 61 TCATCTCTCCCAAACTGCAAGAGACACAGCGCTGACTATATACCGACCAATCATCAAGG 120

QY 946 CTTGCTCTGGCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGATGGTC 1005
 DB 121 ACTTGTCTTGGCTATGCTCTTGGCGGAAACAGACTCGTCAACTGGAACTTTAGAGTGGGC 180

QY 1006 ATTATCTAATTTATTGAATCAACCCAGAGGTGTGAAGAGGCAAGAGATGAATGGACAC 1065
 DB 181 ATTATCTAATTTAGTGAATGACCCAGAGGTGTGCAAGAGGCAAGAGATGAATGGACGC 240

QY 1066 TCAAGTGGGACAGACCGCTGTGTAATGAGTACAGACCTTCCAAACTTCCATATCTTAG 1125
 DB 241 TCAAGTAGGACAGACCGCTGTGTAATGAGTACAGACCTTCCAAACTTCCATATCTCAN 300

QY 1126 GAAGATCATCTTGGACACTTAGGTTGTACCCCGCCGCCCAATTCCTAATCCTCATGT 1185
 DB 301 GAAGATAGTCTTGGAAACACTTAGGTTGTACCCCGCTGCCCAATTCCTAATCCTCATGT 360

QY 1186 GTCTTCAGAGATATTACAAATGAAGGATTCATATCCACGAGACACAAATTTGTGATCAT 1245
 DB 361 GGCTTCAGAGATCATCAATGCAAGGATTCATATCCACGAGACACAAATTTGTGATCAT 420

QY 1246 TAATGTTGGGCGATGACAGAGATCCTCAGTTGTGGAATGATGCCACATGCTTTAAACC 1305
 DB 421 TAATGTTGGGCGATGACAGAGATCCTAAGATATGGAAGATGCGACAGCTTTAAACC 480

QY 1306 TGAGAGTTTGTGTCGAGAGGAGAGAGAAAGTTGGTAGCATTTGGCATGGAAGAG 1365
 DB 481 TGAGAGTTTGTGTCGAGAGGAGAGAGAGAGTTGGTAGCATTTGGTAGGAGAGAG 540

QY 1366 GGCTTCCCGAGAGAACCCATGGCTATGCAAGT 1399
 DB 541 GGCTTCCCGAGAGAACCCATGGCTATGCAAGT 574

RESULT 10
 BM187538
 LOCUS
 DEFINITION
 saj86a06.y1 Gm-cl074 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl074-3467 5' similar to TR:Q9AFX0 Q9AFX0 CYTOCHROME P450
 MONOOXYGENASE ; mRNA sequence.

ACCESSION
 VERSION
 BM187538
 BM187538.1 GI:17518496
 EST
 Glycine max (soybean)
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE
 1 (bases 1 to 561)
 Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
 , A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 , Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 , R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 , R., Waterston, R. and Willson, R.
 Public Soybean EST Project
 Unpublished

TITLE
 JOURNAL
 COMMENT
 Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com

Seq primer: -40RP from Gibco
 High quality sequence stop: 425.

FEATURES

source

1. 561
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl074-3467"
 /tissue_type="seedlings induced for HR (hypersensitive
 response)"
 /dev_stage="9-11 day old"
 /lab_host="DH10B"
 /clone_lib="Gm-cl074"
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from 9-11 day old seedlings that were induced for HR
 (hypersensitive response) by vacuum infiltrating plant
 tissue with Pseudomonas syringae pv. glycinea carrying the
 avrB gene (Genetics 141:1597-1604). Plant tissue (expanded
 unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36,
 and 53 hrs after inoculation and their mRNA pooled equally
 for cDNA construction. The library was prepared using the
 Stratagene pBluescript II SK(+) library construction kit.
 Complementary DNA was synthesized from mRNA using a primer
 consisting of a poly(dT) sequence with an XhoI restriction
 site. EcoRI adaptors were ligated to the blunt-ended cDNA
 fragments followed by XhoI digestion. The cDNA insert is
 protected from XhoI digestion via methylation during first
 strand synthesis. The cDNA fragments were directionally
 cloned into the EcoRI-XhoI restriction site of the
 pBluescript vector. The ligated cDNA fragments were
 transformed into E.coli Electromax DH10B host cells. Plant
 care, inoculations, and library construction were
 performed by Steve Clough (Lila Vodkin lab, University of
 Illinois)."

BASE COUNT 172 a 117 c 130 g 142 t
 ORIGIN

Query Match 24.5%; Score 455.4; DB 12; Length 561;
 Best Local Similarity 88.2%; Pred. No. 1.1e-91;
 Matches 495; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 711 TGTGGAACTCATGGCGTTGGCTAAAGGAGATCACTTCCTTCAGGTGGTCG 770
 DB 1 TGTGCAACTCATGGGTTGGCTAAAGGAGATCACTTACCTTTCCTCAAGGTGGTCG 60

QY 771 ATTTTCAGATGTCGAGAGCGCTTAAGAGTATCAGTAAGAGGTACCATCTTGA 830
 DB 61 ATTTTCAGAGCTGGAGAGAGGTTGAAGATATCAGTAAGAGGTATGATCATTCTGA 120

QY 831 ATAGATCCTTTCATGAGAACCGTGCAGCAATGACCGCAGAAATTCATGATCATC 890
 DB 121 ATAGATCCTTTCATGAGAACCGTGCAGCAATGACCGCAGAAATTCATGATCATC 180

QY 891 TCCTCAAACTGCAAGAGACCCAGCGCTCAGTACTACACTGACCAATCATCAAGGCCCTTG 950
 DB 181 TCCTCAAACTGCAAGAGACACAGCGTCTATTATACCGCAAAATCATCAAGGACTTG 240

QY 951 CTCTGGCCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGAGTGGTCAATTAT 1010
 DB 241 CTCTGGCTATGCTCTTTGGCGGAAACAGACTCGTCAACTGGAACCTTAGAGTGGGCAATTAT 300

QY 1011 CTAATTTATTGAATCACCACAGAGGTGTTGAAGAGGCAAGAGATGAATTTGCACACTCAAG 1070
 DB 301 CTAATTTAGTGAATGACCACAGAGGTGCTGCAAGAGGCAAGAGATGAGTTGACCGCTCAG 360

QY 1071 TGGGACAGACCGCTTGTAAATGAGTCAGACCTTCCAAACTTCCATATCTTAGGAAGA 1130
 |||||
 Db 361 TAGGACAGATCGGCTGTAAATGAGTCAGACCTTCCAAACTTCCATATCTTAGGAAGA 420
 |||||
 QY 1131 TCATCTTGGACACCTTAGTGTGACCCCGGCCCAATCTAATACCTCATGTGTCCT 1190
 |||||
 Db 421 TAGTCTTGAACACTTAGTGTGACCCCGGCCCAATCTAATACCTCATGTGTCCT 480
 |||||
 QY 1191 CAGAAGATATTACAAATGAAGATTCAATATCCACGAGACACAAATGTGATCATTAATG 1250
 |||||
 Db 481 CAGAGACATCATATCGAAGATTCAATGTTCACGAGACACAAATGTGATTAATG 540
 |||||
 QY 1251 GTTGGGCATCGAGAGATC 1271
 |||||
 Db 541 GTTGGGCATCGAAGATC 561
 |||||

RESULT 11

BQ079433 544 bp mRNA linear EST 04-APR-2002
 LOCUS san13f02.y1 Gm-cl084 Glycine max cDNA clone SOYBEAN CLONE ID:
 DEFINITION Gm-cl084-3724 5', similar to SW:CB1E_GLYEC P93147 CYTOCHROME P450
 81E1, mRNA sequence.

ACCESSION

BQ079433

VERSION

BQ079433.1 GI:19934403

KEYWORDS

EST.

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 544)
 Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
 , Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
 , R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
 , R., Waterston, R. and Wilson, R.
 Public Soybean EST Project

TITLE

Unpublished

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com

Trace considered overall poor quality

Seq primer: -40RP from Gibco

High quality sequence stop: 1.

FEATURES

source

1. 544
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl084-3724"
 /tissue_type="Etiolated hypocotyls (Williams 82)"
 /lab_host="DH10B"
 /clone_lib="Gm-cl084"
 /note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; The cDNA library was constructed by M. Bhattacharyya
 from mRNA isolated from etiolated hypocotyls from the
 cultivar Williams 82. Tissue was inoculated with
 Phytophthora soyae race 1 and tissues were harvested 2 and
 4 hours following infection. The library is the pool of
 these two time points. Complementary DNA was synthesized
 from mRNA using a primer consisting of a poly(dT) sequence

with a XhoI restriction site. EcoRI adapters were ligated
 to the blunt-ended cDNA fragments followed by XhoI
 digestion. The cDNA fragments were directionally cloned
 into the EcoRI-XhoI restriction site of the pBluescript
 vector. The ligated cDNA fragments were transformed into
 DH10B host cells (Gibco BRL). This library was constructed
 by M. Bhattacharyya in the laboratory of Dr. Randy
 Shoemaker at Iowa State University."

BASE COUNT 179 a 109 c 111 g 145 t

ORIGIN

Query Match 24.3%; Score 451.2; DB 13; Length 544;
 Best Local Similarity 89.3%; Pred. No. 9.5e-91;
 Matches 486; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 703 GACAGAAATGTTGGAACTCATGGCTTGGCTTAAAGGAGATCACTTCCCTTCTCTCAG 762
 |||||
 Db 1 GACACACATGCTAGAACTCATGGGGTGGCTAATAGGAGATCACTTACCTTCTCTAAG 60
 |||||
 QY 763 GTGGTTCGATTTTCAGAAATGTGGAGAGCGCTTAAAGAGATATCAGTAAGAGGTACGATTC 822
 |||||
 Db 61 GAGGATCGATTTTCAGAAATGTGGAGAGCGCTTGAAGAGATATTAATAAAGGTATGATAC 120
 |||||
 QY 823 CATCTTGAATAGATCCTTCATGACAAACCGTCCAGCAATGACCGCAGCAATTCATGAT 882
 |||||
 Db 121 CATCTTGAATGAGATCATTTGATGAAAACCGTATTAAAGAGGACCGCGAGATTTCCATGAT 180
 |||||
 QY 883 CGATCATCTCTCTCAAACTGCAAGAGACCCAGCTCAGTACTACACTGACCAAAATCATCAA 942
 |||||
 Db 181 TGATCATCTCTCTCAAACTGCAAGAGACTCAGCTGAGTACTACTACCTGACCAAAATCATCAA 240
 |||||
 QY 943 AGGCTTGTCTGTGGCCATGCTTTTGTGGAACTGACTCATCACTGGGACTTTAGAGTG 1002
 |||||
 Db 241 AGGCTTGTCTGTGGCCATGCTTTTGTGGGAACTGACTCATCAACGGGAACCTTAGAGTG 300
 |||||
 QY 1003 GTCAATATCTAATTTTGAATCAACCCAGAGCTGTGGAAGAGCAAGAGATGAATTGGA 1062
 |||||
 Db 301 GTGCTATCTAATTTTGAATCAACCCAGAGCTGTGGAAGAGCAAGAGATGAATTGGA 360
 |||||
 QY 1063 CACTCAAGTGGGACAAAGCGCTTGTAAATCAGTCAGACCTTCCAAAACTTCCATATCT 1122
 |||||
 Db 361 CACTCAAGTGGGACAAAGCGCTTGTAAATCAGTCAGACCTTCCAAAACTTCCATATCT 420
 |||||
 QY 1123 TAGGAAGATCATCTCTTGAGACACTTAGGTGTGACCCCGGCCCAATTTAATACCTCA 1182
 |||||
 Db 421 TAAAAGATCATGCTTTGAGACACTTAAGTTGTATCTCCAGCTCCAATTTCTAATACCTCA 480
 |||||
 QY 1183 TGTGCTTCCAGAAATATTACAAATTAAGAGATTAATATCCACGAGACACAAATTTGAT 1242
 |||||
 Db 481 TGTGCTTCCAGAAATATTACAAATTAAGAGATTAATATCCACGAGACACAAATTTGAT 540
 |||||
 QY 1243 CAT 1246
 |||||
 Db 541 CAT 544
 |||||

RESULT 12

BQ453555

LOCUS

DEFINITION

BQ453555 564 bp mRNA linear EST 29-MAY-2002
 sao83f08.y1 Gm-cl081 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl081-2559 5', similar to TR:Q9AFX0 Q9AFX0 CYTOCHROME P450
 MONOOXYGENASE ; mRNA sequence.

ACCESSION

BQ453555

VERSION

BQ453555.1 GI:21256667

KEYWORDS

EST.

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max

REFERENCE

1 (bases 1 to 564)

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna

FEATURES	source
Location/Qualifiers	
1. .564	
/organism="Glycine max"	
/mol_type="mRNA"	
/db_xref="taxon:3847"	
/clone="SOYBEAN CLONE ID: Gm-cl074-9600"	
/issue_type="seedlings induced for HR (hypersensitive response)"	
/dev_stages="9-11 day old"	
/lab_host="DH10B"	
/clone_lib="Gm-cl074"	
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with <i>Pseudomonas syringae</i> pv. <i>glyciniae</i> carrying the avB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into <i>E. coli</i> ElectroMax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Uila Vodkin lab, University of Illinois)." 160 a 148 c 136 g 120 t	
BASE COUNT	160 a 148 c 136 g 120 t
ORIGIN	
Query Match	22.9%; Score 426; DB 13; Length 564;
Best Local Similarity	86.7%; Pred. No. 4.5e-85;
Matches	494; Conservative 0; Mismatches 70; Indels 6; Gaps 2;
QY	215 CGTTTCTTCAAGCATGTCGAAACAGTACCGACGTTGGTTTCCCTCTGTTGGTTCA 274
DB	1 CGTTTCTTCAAGCATGTCGAAACAGTATGGAACATCGTTTCCCTCTGTTGGTTCA 60
QY	275 CGTCTCGCGTGTGTCATCTCTTCAACAGCATACCAAGATGTTCCACCAACACGAC 334
DB	61 CGTCTCGCGTGTGTCATCTCTTCAACAGCATACCAAGATGTTCCACCAACACGAC 120
QY	335 GTTGCTTTGGCCAAACCGGTGTCCTTCTCTCGGAAATACATCTTCTCAACACAC 394
DB	121 GTTGCTTTGGCCAAACCGGTGTCCTTCTCTCGGAAATACATCTTCTCAACACAC 180
QY	395 ACCGTAGGTCCTGCTCCACCGGAGACTGGCGCACTCCGCGCATCACGCCCTG 454
DB	181 ACCGTAGGTCCTGCTCCACCGGAGACTGGCGCACTCCGCGCATCACGCCCTG 240
QY	455 GAGTCTCTTCAACGAGCGTGCACCTCTTCTCGGAATCCGAGGCGAGACGAAG 514
DB	241 GAGTCTCTTCAACGAGCGTGCACCTCTTCTCGGAATCCGAGGCGAGACGAAG 300
QY	515 CGTCTGATGAGGTTGGTGTGGCCAGAACTCGAACGAGGAGAGTTTTCGCGAGTG 574
DB	301 AGGTTGGTACAGAGT---TGCTGACAGAACTC---CAAGGAGGGTTTTCGCGAGTG 354
QY	575 GAGATTAGTTCGATGTTCAACGACTTAATTCACACATATATGAGGATATATCGGG 634
DB	355 GAGATTAGTTCGATGTTCAACGACTTAATTCACACATATATGAGGATATATCGGG 414
QY	635 AAGAGGTTTTACGAGGAGGAGTGAGATGAAGAAGCTTGAGGAGGCGAGGAGTTTCA 694
DB	415 AAGAGGTTTTACGAGGAGGAGTGAGATGAAGAAGCTTGAGGAGGCGAGGAGTTTCA 474
QY	695 GAGACTGTGACGAATATGTTGGAACCTCATSGGCTTGGGTTAAACAGGGAGATCATTTGCT 754

Query Match 22.4%; Score 417; DB 12; Length 425;
 Best Local Similarity 98.8%; Pred. No. 4.4e-83;
 Matches 420; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1128 AGATCATCTTGAGACACCTTAGGTGTGACCCCGCCGCCCAATCTTAATACCTCATGTCT 1187
 DB 1 AGATCATCTTGAGACACCTTAGGTGTGACCCCGCCGCCCAATCTTAATACCTCATGTCT 60

QY 1188 CTTGAGAAGATATTCAATTTGAAGGATTCAATATCCACGAGACACAAATTTGATCATATTA 1247
 DB 61 CTTGAGAAGATATTCAATTTGAAGGATTCAATATCCACGAGACACAAATTTGATCATATTA 120

QY 1248 ATGTTTGGGCATGCGAGAGATCTCAGTTGTGGAAATGATGCCACATGCTTTAAACCTG 1307
 DB 121 ATGTTTGGGCATGCGAGAGATCTCAGTTGTGGAAATGATGCCACATGCTTTAAACCTG 180

QY 1308 AGAGGTTTGTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1367
 DB 181 AGAGGTTTGTATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240

QY 1368 CTTGCCCGAGAGAACCCATGGCTATGCAAGTGTGAGCTTTTACTTTGGGATTTGTTGATTC 1427
 DB 241 CTTGCCCGAGAGAACCCATGGCTATGCAAGTGTGAGCTTTTACTTTGGGATTTGTTGATTC 300

QY 1428 AATGTTTGTGCTGGAACAGGTAAGTGAGAGAAAGCTTGATATGACAGAGAACAAATTGA 1487
 DB 301 AATGTTTGTGCTGGAACAGGTAAGTGAGAGAAAGCTTGATATGACAGAGAACAAATTGA 360

QY 1488 TCACCTTGTCAAGGTTAAATTCATTTGGAGCCATGTCAGGCTCGCCCATCTGCCACTA 1547
 DB 361 TCACCTTGTCAAGGTTAAATTCATTTGGAGCCATGTCAGGCTCGCCCATCTGCCACTA 420

QY 1548 AAATT 1552
 DB 421 AAATT 425

RESULT 15
 BI426262
 LOCUS saf09b06.y3 Gm-c1076 Glycine max cDNA clone GENOME SYSTEMS CLONE
 DEFINITION ID: Gm-c1076-84.5, similar to FR:Q9ZRW6 Q9ZRW6 CYTOCHROME P450 ,
 mRNA sequence.

ACCESSION
 VERSION BI426262.1 GI:15203494
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 426)
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
 ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 ,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project

TITLE Unpublished
 JOURNAL
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Putative full length read
 vector to vector length is This clone is available through: ResGen,
 Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801
 For further information call: (800)-533-4363 or contact via email:

ccu@resgen.com

High quality sequence stop: 398.

Location/Qualifiers

1. 426

source

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1076-84"

/tissue_type="wounded cotyledons"

/dev_stage="11 day old seedlings"

/lab_host="DH10B"

/clone_lib="Gm-c1076"

/note="Vector: pBluescript II SK+; Site1: EcoRI; Site2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from 11 day old seedlings treated with that were treated
 with 2 ugs/ml of a crude glucan elicitor preparation
 isolated from the mycelial walls of Phytophthora sojae.

The library was prepared using the Stratagene pBluescript
 II SK(+) library construction kit. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) sequence with an XhoI restriction site. EcoRI
 adaptors were ligated to the blunt-ended cDNA fragments
 followed by XhoI digestion. The cDNA fragments were
 directionally cloned into the EcoRI-XhoI restriction site
 of the pBluescript vector. The ligated cDNA fragments were
 transformed into E.coli ElectroMax DH10B host cells. Plant
 material was provided by Michael G. Hahn (Complex
 Carbohydrate Research Center, University of Georgia) and
 the library was constructed by Anu Khanna (Lila Vodkin lab
 at the University of Illinois)."

BASE COUNT 140 a 58 c 79 g 149 t

ORIGIN

Query Match 22.1%; Score 410.8; DB 12; Length 426;

Best Local Similarity 98.3%; Pred. No. 1.1e-81;

Matches 415; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1408 TACTTTGGGATTTGTTGATTCATGTTTGGCTGGAACAGGTAAGTGAGGAAAGCTTGA 1467

DB 5 TTCTGTTTCGAGGGTTGATTCATGTTTGGCTGGAACAGGTAAGTGAGGAAAGCTTGA 64

QY 1468 TATGACAGAGAACAAATTCGATCACCCTTGTCAAGGTTAATTCATTGGAGGCCATGTGCAA 1527

DB 65 TATGACAGAGAACAAATTCGATCACCCTTGTCAAGGTTAATTCATTGGAGGCCATGTGCAA 124

QY 1528 GGCTCGCCCATCTGCCACTAAATTTGAATTTAATTAATAGTATTTTATTTGGTAA 1587

DB 125 GGCTCGCCCATCTGCCACTAAATTTGAATTTAATTAATAGTATTTTATTTGGTAA 184

QY 1588 ACTTGGGTGATTCAGAAATCTAATCTTAATTTTAGTGTGAAGAGTGGTATCATATA 1647

DB 185 ACTTGGGTGATTCAGAAATCTAATCTTAATTTTAGTGTGAAGAGTGGTATCATATA 244

QY 1648 TACATTTCAAATTAATTAATCTTTGTCCAAATTCATCCATGCAACATATATGCAATT 1707

DB 245 TACATTTCAAATTAATTAATCTTTGTCCAAATTCATCCATGCAACATATATGCAATT 304

QY 1708 GACATCTAGAGAGAAATAGATATAAGAAATTTTATATTTTATCTCTCTTTATCTT 1767

DB 305 GACATCTAGAGAGAAATAGATATAAGAAATTTTATATTTTATCTCTCTTTATCTT 364

QY 1768 ATGTCTCAAGGCCCATTTAGAAATTTGGGTGAGCATTAACATATATCAATATTTGTATCCG 1827

DB 365 ATGTCTCAAGGCCCATTTAGAAATTTGGGTGAGCATTAACATATATCAATATTTGTATCCG 424

QY 1828 CC 1829

DB 425 CC 426

Search completed: October 21, 2003, 05:47:14

Job time : 3611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 01:48:19 ; Search time 500 Seconds
(without alignments)
10036.511 Million cell updates/sec

Title: US-09-868-546A-1
Perfect score: 1859
Sequence: 1 gaaacacgacagacagca.....aataaattcttacttctc 1859

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1859	100.0	1859	21	AAA29326 Soybean isoflavone
2	1122	60.4	1698	21	AAA29327 Soybean isoflavone
3	396.2	21.3	1497	21	AAC50087 Arabidopsis thalia
4	393.8	21.2	1488	24	ABZ13383 Arabidopsis thalia
5	386.6	20.8	1494	21	AAC43264 Arabidopsis thalia
6	386.6	20.8	1494	24	ABZ14473 Arabidopsis thalia
7	386.6	20.8	1494	25	ABZ42120 Arabidopsis thalia
8	386.6	20.8	1654	21	AAC50097 Arabidopsis thalia

9	384.6	20.7	1655	21	AAC47389 Arabidopsis thalia
10	384.6	20.7	1656	21	AAC37476 Arabidopsis thalia
11	383.8	20.6	1479	24	ABZ14367 Arabidopsis thalia
12	383.8	20.6	1545	21	AAC42545 Arabidopsis thalia
13	383.2	20.6	1503	24	ABZ13668 Arabidopsis thalia
14	381	20.5	1653	21	AAC41374 Arabidopsis thalia
15	377	20.3	1503	24	ABZ12256 Arabidopsis thalia
16	377	20.3	1664	21	AAC47894 Arabidopsis thalia
17	374.2	20.1	1665	21	AAC50095 Arabidopsis thalia
18	361.8	19.5	1509	24	ABZ13030 Arabidopsis thalia
19	361.8	19.5	1519	21	AAC47053 Arabidopsis thalia
20	354	19.0	1576	21	AAC47416 Arabidopsis thalia
21	354	19.0	1578	21	AAC35968 Arabidopsis thalia
22	351.4	18.9	1674	20	AAAS8406 Jerusalem artichok
23	350.6	18.9	1557	24	ABZ14154 Arabidopsis thalia
24	330.4	17.8	1548	24	ABZ13736 Arabidopsis thalia
25	328.4	17.7	1698	19	AAV63107 Euphorbia lagascae
26	244.4	13.1	843	21	AAA29328 Soybean isoflavone
27	202.8	10.9	722	24	ABQ65567 Arabidopsis thalia
28	182.2	9.8	447	21	AAA68128 Eucalyptus grandis
29	181.6	9.8	1545	22	AAAS0178 Taxus cuspidata ox
30	168.6	9.1	475	24	ABL93297 Arabidopsis thalia
31	159.4	8.6	1530	25	AAAD4749 Tobacco cytochrome
32	158.2	8.5	1488	24	ABL61191 P. sativum ddfw1 c
33	143.4	7.7	494	21	AAA68129 Eucalyptus grandis
34	139.8	7.5	1770	21	AAA88000 Perilla frutescens
35	139.4	7.5	495	25	ABX61012 Arabidopsis thalia
36	139	7.3	1557	24	ABZ14864 Arabidopsis thalia
37	135.6	7.3	1789	18	AAAT94655 Petunia flavonoid
38	131.8	7.1	1704	20	AAAG60781 Soybean cytochrome
39	130	7.0	1784	21	AAC34631 Arabidopsis thalia
40	129.6	7.0	1770	22	AAF30281 Perilla flavone sy
41	129.4	7.0	1650	21	AAC42321 Arabidopsis thalia
42	128.6	6.9	1812	14	AAQ34489 Insert in pCGP176
43	128.6	6.9	1812	14	AAQ49417 Cytochrome P450 ho
44	128.6	6.9	1824	14	AAQ47842 Flavonoid-3',5'-hy
45	127.2	6.8	522	19	AAV23837 Plant C4H enzyme D

ALIGNMENTS

RESULT 1
AAA29326
ID AAA29326 standard; cDNA; 1859 BP.
XX
AC AAA29326;
XX
DT 26-SEP-2000 (first entry)
XX
DE Soybean isoflavone-2-hydroxylase coding sequence.
XX
KW Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;
KW flower colour; pollen tube; feeding deterrent; UV irradiation; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
CDS 59..1561
FT /*tag= a
FT /product= Isoflavone-2-hydroxylase
XX
PN WO200037652-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US30337.
XX
PR 21-DEC-1998; 98US-0113190.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;


```
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.3%; Score 396.2; DB 21; Length 1497;
Best Local Similarity 57.0%; Pred. No. 7.3e-83;
Matches 838; Conservative 0; Mismatches 603; Indels 30; Gaps 5;

QY 90 TTTTCTCTGTTCTATTCTCTCGCGGGAAGTTTGTTCCTCCAAAGCAGAAAATTTGAGAAACA 149
DB 29 TTCTCTTCTGTTGTTCTCTCTCTCATTTACTTAATTTGAAAATCTCAGCGAAGCCAAATC 88

QY 150 TACCACCAAGTCTCC---TCCTCTTCCCATTAATAGGAAACCTTAACCTCTCGAAGAGC 206
DB 89 TACCTCCGAGTCCGGCATGGTCATTACCGGTGATTGGTCATCTCGCCCTTCTCAAAACCAC 148

QY 207 CAATCCACGGTTCTTCCACGCATGTCGAACAGTACGGCAACG-----TGGTTTCCC 260
DB 149 CGAATTCATCGCACATTTCTCTCCCTCTCTCAATCCCTAAACAATGCTCCGATCTTCTCCC 208

QY 261 TCTGTTTCGTTTCAGTCTGCGCGTGTGTCATCTCTCTCCACAGCATACCAAGAAATGCT 320
DB 209 TCCGACTCGTAAACCGACTCGTTTTTCGTGAATCTGTCACTCGATCGCGGAGGATGTT 268

QY 321 TCACCAAAACAGACGCTTGCCTTGGCCAAACCGGCTACCTTCTCTCTCGGAAAATACATCT 380
DB 269 TCACCAAAACAGACGCTGCTACTGGGAAACAGACACCAAACTTCATCTCTCGCTAAACACGTTG 328

QY 381 TCTACAAACAGACACCGTAGCTCTGCTCCACGGGAGACACTGGCGCAACCTCCGCC 440
DB 329 CGTACGATTACAAACCATGATCGCAGCTTCTTACGGTGACCACTGGCGGTAACTCCGCC 388

QY 441 GCATCACCGCCCTGGACGCTCTCTCCAGCAGCGGTCACCTCTCTTCTCCGGAATCCGGA 500
DB 389 GCATCGGCTCCGTCGAGATATTCTCCAAATCACCCTCAATAGCTTTTGTCTATTCTGTA 448

QY 501 GCGACGAGAGGAAGGCTCTGATGCGAGAGTTGGTGTGCGGCCAAGAACTCGAAGGAGGAAG 560
DB 449 AAGACGAGATCCGACGACTTGTGTTTCGTTCTTCTCGGAAC-----TTTTCACAAG 499

QY 561 AGTTTTCGCGGAGTGCAGATTAGTTTCGATGTTTCAAACGACTTAATTACACACATAATCA 620
DB 500 AGTTTGTGAAAGTGGATATGAATCAATGTTTATCTGACTTAACATTCACCAACATTTTAA 559

QY 621 GGATCATATCGGGAAGAGGTTTTCACGGAGAGGAGGTGAGATGAAGAACGTTTGAGGAAG 680
DB 560 GAATGTTGCGCGAANAACGTTTACTACGGAGACGTTGTGAGG-----ATGATCCGGAGG 613
```


Db 467 GGCTCATACACATCTCTAGAGACT-----CCTTGACGGATTGTTGAAGTGG 517
Qy 576 AGATTAGTTGATGTTCAACGACTTAACATCAACACATAATAGAGATGATATCGGGA 635
Db 518 AGTTGAAATCGTTATTAAACCACTTGGCATTCACCAACATCATCATGATGGTAGCGGAA 577
Qy 636 AGAGGTTTTACGAGAGGAGAGTGAGATGAAGAACGTTGAGGAGCGAGGAGTTGAGAG 695
Db 578 AAGCATATTACGG-----TACCGGTACAGAGACAATGATGAAGCAAGCAAGCTCGTAGGG 631
Qy 696 AGACTGTGACAGAAATGTTGGAATCATGGCTTGGCTTAACAAAGGGAGATCACTGCCCTT 755
Db 632 AACTTAAGCGGAGATAATGGCCGGCGCGTCTCGAAATTTGGCTGATATATCTTCCGT 691
Qy 756 TCCTCAGGTGGTTTCGATTTTCAGAAATGTGAGAGAGCGCTTAAAGATATCAGTAAGAGT 815
Db 692 CCATAAATTTGGTCT-----ACAACTTTGAGAACCCAGACCAAAATCTTGGGGAATCGAC 745
Qy 816 ACGATTCCATCTTGAATAGATCCCTTCATGAGAACCGTGCAGCAATGACCGCCAGATT 875
Db 746 TCGATAGAGTCTTGCAGAAATCGTTGATGAGAAACGTCGAGAGAAAGAAAGGGTCAAA 805
Qy 876 CCATGATCGATCATCTCCTCAAACTGCAAGAGACCCAGCTCAGTACTACACTGACCAAA 935
Db 806 CTTTGATCGATCATTTGCTTCTTTCAGAAACCGAACCCGAGTACTATCTGATGTCA 865
Qy 936 TCATCAAGGCTTGTCTGGCGATGCTTTTGGTGGAACTGACTCATCACTGGGACTT 995
Db 866 TCATCAAGGAAATCATACTGGCTTGTCTTGGGGGACAGATACATCTTCAGTGAAGT 925
Qy 996 TAGATGGTCATTATCTAATTTATGAAATCACCAGAGAGTGTGAAAGAGGCGCAAGAGATG 1055
Db 926 TGAATGGGCAATGTCAAAATTTGCTGAACCATCCAGAGATACCTTGAGAAAGCGAGAGCG 985
Qy 1056 AATTGGACACTCAAGTGGGCAACAGCGCTTGTAAATGAGTCAGACCTTCCAAACTTC 1115
Db 986 AGATCGATGATAAATCGGTTGACCGGTTAGTGAAGAAATCAGATATTTGTAATCTCC 1045
Qy 1116 CATATCTTAGGAAGATCATCTTTGAGACACTTAGTGTGTACCCCGGCCCAATTTCTAA 1175
Db 1046 ATTATCTCAAAACATTTGTGTCAGAAACATTACGTTGTATCTCGCGTTCCACTACTAC 1105
Qy 1176 TACCTCATGTGCTTCAGAGATATTACAAATTCAGGATTCATATCCACGAGACACAA 1235
Db 1106 TCCTCATTTCTCATCGGATGAATGTAAGTGGCGGCTACGATATGCCACGTCGCACGT 1165
Qy 1236 TTGTGATCATTAATGTTGGGCGATCGAGAGATCTCTCAGTTGTGGAATGATGCCACAT 1295
Db 1166 TGTATTAAACAGCTATGGCGGATGATAGATCCAGTTTTATGGAGAGCCAGAGA 1225
Qy 1296 GCTTTAAACCTGAGAGTTGATGTGGAAGGAGAGAGAAAGTTGGTAGCATTTGGCA 1355
Db 1226 GGTTCGAAGCGGAGAGGTTTCAGAAAGAGAGAGGCTTCGAAAGCTAATGCCGTTTGGGA 1285
Qy 1356 TGGGAGAGGCGCTGCCAGGAGAACCCATGCTATGCAAACTGTCAGCTTACCTTTTGG 1415
Db 1286 TGGGAGAGAGCTTGTCTGGAGCTGAGCTTGGGAGCGGTTAGTGAGCCTTGTCTCTT 1345
Qy 1416 GATTGTTGATTTCAATGTTTTCAGTGGAAACGAGTAAAGTGAAGAAAGCTTGTATGACAG 1475
Db 1346 GGTGCTTGTATTCAGTCTTTCAGTGGGAGAGAGTTGTTGAGAGAACTTGTGACATGACTG 1405
Qy 1476 AGAACAAATGGATCACTCTCAAGGTTAATTCATTTGAGGCGCATGTGCAAGGCTCGCC 1535
Db 1406 AAGGCGAAGGATCACTATGCTTAAAGCTACTCCGTTGCGAGCTATGTGCAAGGCAAGT 1465
Qy 1536 CACTTGGCACTAAA 1550
Db 1466 CCATTGTTGGTAA 1480

ID AAC43264 standard; DNA; 1494 BP.
XX
AC AAC43264;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38633.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

```
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.8%; Score 386.6; DB 21; Length 1494;
Best Local Similarity 57.0%; Pred. No. 1.3e-80;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

Qy 90 TTTTCTGGTTCTATTCTCTCGGCGTCAAGTTTGTTCCTTCCAAAGCAGAAAAATTCAGAAACA 149
    |||||
Db 29 TTCTCTGGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 88
    |||||
Qy 150 TACCACCAAGTCTCTCC---TCCTTTTCCCATATAGAAACCTTAACCTCTCTCGAACAGC 206
    |||||
Db 89 TACCTCCGAGTCCGGCATGTGTCATTACCGGTGATTGTCATCTCCGCTTCTCAAAACCAC 148
    |||||
Qy 207 CAATCCACGTTTCTTCCACAGCATCTCGAAACAGTACGCCAAGC-----TGGTTTCCC 260
    |||||
Db 149 CGATTTCATCGCAATTCCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 208
    |||||
Qy 261 TCTGGTTTCGTTTCAACGCTCTGGCGGTGTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 320
    |||||
```

Db	209	TCCGACTCGGTAAACCGGACTCGTTTTCGTGAACACTCGTCACACTCGATCCGCGAGGAATGTT	268
Qy	321	TCACCAAAACGAGCTGGTTCCTTGCCCAACCGGCTACCTTCTCTCTCGGGAATAATACATCT	380
Db	269	TCACCAAAACGAGCTCGTACTGCGGAAACGACCAAACTTCATCTCGCTAAACAGTTG	328
Qy	381	TCTAACAAACACCAACCGTAGGCTCTGCTCCCAACGGGAGACACTGGCGCAACCTCGGCC	440
Db	329	CGTAGGATTAACAAACCATGATCGAGCTTCTTACGGTGACCACTGGCGTAAACCTCGGCC	388
Qy	441	GCATCACCGCCCTCGGACGCTCTCTCCACGAGCGCGTCCACTCTCTCTCGGAATCCGA	500
Db	389	GCATCGGCTCGTCCGAGATATCTCCCAATCACCGTCTCAATAGCTTTTGTCTATTCTGTA	448
Qy	501	CGCACGAGACGAACCGTCTGATGCAGAGGTTGGTGTCTGGCCAAAGAACTCGAAACGAGGAAG	560
Db	449	AAGACGAGATCCGAGGACTTGTGTTCTCTTCGGAAC-----TTTTCAAG	499
Qy	561	AGTTTGGCGGAGTGGAGATATGTTTCGATGTTCAACGACTTAACATTAACAACAATATGA	620
Db	500	AGTTTGTGAAGTGGATATGAATCAATGTTATCTGACTTAACATTCAACAACATTTTAA	559
Qy	621	GGATCATATCGGGAAAGGTTTTCACGAGGAGAGTGAAGTGAAGAAAGTTGAGGAAG	680
Db	560	GAATGGTGGCGGAAACAGTTACTACCGGACGCGTGTG-----AGGATGATCCGGAGG	613
Qy	681	CGAGGGAGTTCAGAGAGACTGTGACAGAAATGTTTGGAACTCATGGGCTTCGGCTAAACAAGG	740
Db	614	CTAAACGTGTCGGGACGCTTATACGGATGTGGTGGCTTGTGTGGTGTGGAACCGCTG	673
Qy	741	GAGATCACTTGCCTTTCTCCTCAGGTGGTTTCGATTTTCAGAAATGTGGGAAGCGCTTAAAGA	800
Db	674	TTGATTACTTACCGGTTTTCGGTGG-----TTTCAGATTACGACACCGGGTTAAGA	727
Qy	801	GTATCAGTAGAGGTAGTATCCATCTTTGAATAAGATCCTTTCATGAGAACCGTGCCACGA	860
Db	728	AGTTGGCGGGTAGGCTCGACGAGTCTTTCGCAAGGATTTGGTTGATGAGAAACGGAAGCTA	787
Qy	861	ATGACCGCCAGAAATCCATGATCCATCATCTCTCAAACTGCAAGAGACCCAGCCTCAGT	920
Db	788	AGGAGAAAGGAAACACTATGATCGATCATCTTCTTACTCTTGCAAGAAATCCAACCGGATT	847
Qy	921	ACTACACTGACCAAATCATCAAGGCCCTTGCTCTGGCCATGCTTTTGGTGGAACTGACT	980
Db	848	ACTTCAGGATCGTATCATTAAGAGAAACATGCTCGCTTGATCTAGCAGGAGCCGACA	907
Qy	981	CATCAACTGGGACTTTAGAGTGGTCTATTACTATTTATTGAATCACCCGAGGTGTGA	1040
Db	908	CATCAGCGGTTACGTTAGAAATGGGCATTTGTGCAACGTTGTTGAACCATCCGGATGTATTGA	967
Qy	1041	AGAAGGCAAGAGATGAATTGACACTCAAGTGGCAACAAGCCGCTTGTTAAATGAGTCAG	1100
Db	968	ACAAGGCGAGAGATGAATCGATAGAAGATAGGTTTAGACAGGCTTATGGATGAATCAG	1027
Qy	1101	ACCTTCCAAAACCTTCCATATCTTAGGAAGATCATCTCTTGAGACACTTAGGTTGTACCCCC	1160
Db	1028	ATACTCAACCTGCCCTTATCTCAAAACATTTGTGCTGAAACGTTTCGCGCTTATTCCTG	1087
Qy	1161	CGGCCCCAATTCATAATACCTCATGTGTCTTCAGAAAGATATTACAAATTTGAAGGATTCAATA	1220
Db	1088	CGGCTCCCATGCTTCTTCTCAGCGTTCGCTCGGAAGATTGTAAAGTTTGCAGGATACGATA	1147
Qy	1221	TCCACACGAGACACAATTTGTGATCAATTAATGGTTGGGCGCATGCACAGAGATCCCTCAGTTGT	1280
Db	1148	TGCCGCTGGCACGATACTTATTGACCAATGTGTGGGCTTATACACAGAGATCCTCAGCTAT	1207
Qy	1281	GGAAATGATGCCACATGCTTTTAAACCTCAGAGGTTTGTATGTGGAAGGAGAGAGAAAGT	1340
Db	1208	GGGATGATCCAATGAGCTTCAAGCCAGAGAGTTTGAGAAGAAGAGAGAGCTCAGAGAC	1267
Qy	1341	TGGTAGCAATTTGGCATGGGAAGAGGGCTTCCCGAGGAGAACCCATCGCTATGCAAGTG	1400
Db	1268	TAATGCCGCTTTGGGTAGGAAGAGGGCGTCTCTGTTCTTGAGACTGGCTCATCGGCTTA	1327

1401 TCAGCTTTACTCTTTGGGATGTTCTCATTCATGTTTTGCCTCGGAACGAGTAAGTGCAGGAA 1460

1328 TAAACCTGACTCTTGGATCATTTGATCTCAGTGTGTTGGAATGGGA---GAAGATTGGAGAG 1384

1461 AGCTTGATATGACAGAGAGACAATTTGGATCACCTTGTCAAGGTTAATTCATTCGAGGCCA 1520

1385 AAGTGGATATGAGTGAAGGCAAGAGGTGTTACAATGCTTAAAGCCAAGCCTTTTGGAAAGCCA 1444

1521 TGTGCAAGGCTCGCCCACTTGCCCACTAAAAT 1551

1445 TGTGCAGACACGTCCTCCTCTGTTGTTAAAAT 1475

RESULT 6	
ABZ14473	
ID	ABZ14473 standard; DNA; 1494 BP.
XX	
AC	ABZ14473;
XX	
DT	21-JAN-2003 (first entry)
XX	
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 2278.
XX	
KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200216655-A2.
XX	
PD	28-FEB-2002.
XX	
PF	24-AUG-2001; 2001WO-US26685.
XX	
PR	24-AUG-2000; 2000US-227866P.
PR	26-JUN-2001; 2001US-264647P.
PR	22-JUN-2001; 2001US-300111P.
XX	
PA	(SCRI) SCRIPPS RES INST.
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Harper JF, Krepis J, Wang X, Zhu T;
XX	
DR	WPI; 2002-304127/34.

100

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132207.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148665.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 08-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

```

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      20.8%; Score 386.6; DB 21; Length 1654;
Best Local Similarity 57.0%; Pred. No. 1.3e-80;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

QY 90 TTTTCTCGTTCATCTCTCGGCGTGAAGTTTGTTCCTCAAGCAGAGAAATTCAGAAACA 149
DB 46 TTCTCTTCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 105

QY 150 TACCACAGGTCTCC---TCCTCTTCCCATATAGGAACCTTAACCTCTCGAACAGC 206
DB 106 TACCTCGAGTCGGCATGTGTCATTACCGGTGATGTGTCATCTCCGCTTCTCAACCCAC 165

QY 207 CAATCCACCGTTTCTTCCACGCATGTCGAAACAGATCAGGCAACG-----TGGTTTCCC 260
DB 166 CGATTATCGACATCTCTCTCCCTCTCTCAATCCCTTAACCAATGCTCCGATCTTCTCCC 225

QY 261 TCTGGTTCGGTTCACGTCGCGGTGTGTCATCTCTCTCAACAGCATACCAAGATGCT 320
DB 226 TCCGACTCGGTAACCGACTCGTTTTCGTGAATCGTCACACTCGATCGCGGAGGAATGTT 285

QY 321 TCACCAACACGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATACATCT 380
DB 286 TCACCAAAAGAGCTGCTACTGGCGAAACAGACCAACTTCATCTCTCGCTAAACAGTTG 345

QY 381 TCTAACAAACACACCGTAGGCTCTGCTCCACCGCGAGCACTGGCGCAACCTCGCC 440
DB 346 CGTACGATTACAAACCATGATCGAGCTTCTTACGGTGACCACTGGCGTAAACCTCGCC 405

QY 441 GCATCAGCGCTCGAGCTCTCTCCAGCAGCGCTCCACTCTTCTCGGATCCGGA 500
DB 406 GCATCGCTCGTGGAGATATCTCCAAATCACCGTCTCAATAGCTTTTGTCTATTGCTGA 465

QY 501 CGCAGCAGACGAACGCTCTGATGCAGAGGTTGTGCTGGCCAAAGAACTCGAACGAGGAAG 560
DB 466 AAGACGAGATCCGACGACTTGTGTGCTTCTCTCGAAC-----TTTTTCAAG 516

QY 561 AGTTTCGCGAGTGGAGATTAGTTTCGATGTTCAACAGCTTAACATTAACAAACATATGA 620
DB 517 AGTTTGTGAAGTGGATATGAATCAATGTTATCTGACTTAAACATTCACCAACATTTAA 576

QY 621 GGATGATATCGGGAAGAGTTTTCGAGAGGAGATGAGATCAAGAACTTGAGGAAG 680
DB 577 GAATGTTGGCGGGAACAGTTTACTACCGAGACGCGTGTG-----AGGATGATCCCGAGG 630

QY 681 CGAGGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGGCTTAAACAAG 740
DB 631 CTAAACGTGTCGCGACCTTATACCGGATGTGTGGCTTGTGCTGGTGGAAACGCTG 690

```

```

QY 741 GAGATCACTTGCCTTTCCTCAGGTGGTTCGATTTTCAGAAATGTGAGAAAGCCCTTAAAGA 800
DB 691 TTGATTACTTACCGGTTTTGCGGTTGG-----TTTCAGATTACGAGACACGGGTTAAGA 744

QY 801 GTATCAGTAAGAGGTACGATTCCTTGAATAAGATCCTTCATGAGAACCGTGGCCAGCA 860
DB 745 AGTTGGCGGTAGGCTCGACGAGTTCTTGAAGGATTTGGTTGATGAGAAACGAGAAGCTA 804

QY 861 ATGACCGCCAGAAATTCCTATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCCTCAGT 920
DB 805 AGGAGAAAGGAAACACTATGATCGATCACTTGTCTACTCTGCAAGATTCCTCAACCGGAT 864

QY 921 ACTACACTGACCAAAATCATCAAGCGCTTGTCTGCGCANGCTTTTTTGTGTGGAACCTGACT 980
DB 865 ACTTCACGGATCGTATCATTAAGGAAACATCTCGCTTTGATATCTAGCAGGACCGACA 924

QY 981 CATCAACTGGGACTTTAGAGTGGTCTATTTATTAATTTGAATCACCCAGAGGTGTTCA 1040
DB 925 CATCAGCGTTACGTTAGAAATGGGCAITGTGCAAGCTGTTGAACCATCCGGATGTATGA 984

QY 1041 AGAAGGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTTGTTAAATGATCAG 1100
DB 985 ACAGCGGAGAGATGAATTCGATGAAGATAGGTTTAGACAGGCTTATGGAATGATCAG 1044

QY 1101 ACCTTCCAAACTTCCATATCTTAGAAGATCATCTCTTGAGACACTTAGGTTGTACCCCC 1160
DB 1045 ATATCTCAAAACCTGCTTATCTCCAAACATTTGTCTGAAACGTTTGCCTTTATCTCTG 1104

QY 1161 CGGCCCCCAATTTCACTACTCATGTCTTCAGAAAGATATTACAATTTGAAGGATTCATA 1220
DB 1105 CGGCTCCCAATGCTTCTCTCACGTTCCTCGGAAGATTTGAAGTTGCAGGATACGATA 1164

QY 1221 TCCCAACGACACAAATTTGTGATCAATTAATGTTGGGCGATGCAGAGAGATCCTCAGTTGT 1280
DB 1165 TCGCGGTGGCACGATACTATTGACCAATGTGTGGCTATACACAGAGATCCTCAGCTAT 1224

QY 1281 GGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTATGTGGAAGAGAGAGAAAGT 1340
DB 1225 GGGATGATCCATGAGCTTCAAGCCAGAGAGGTTTGAAGAAAGAGAGAGAGCTCAGAAGC 1284

QY 1341 TGGTAGCATTTGGCATGGGAAGCGGCTTCCAGGAGAAACCCATGGCTATGCCAAAGTG 1400
DB 1285 TAATGCGGTTGGGTTAGGAAGAGGCGGTGCTGGTCTCGACTGGCTCATCGGCTTA 1344

QY 1401 TCAGCTTTTACTTTGGGATTTGTTGATTCATGTTTTCATCTGGAACGAGTAAGTGAGGAAA 1460
DB 1345 TAAACCTGACTCTTGGATCATTTGATTCAGTGTGGAATGGGA---GAAGATTGGAGAAG 1401

QY 1461 AGCTTGTATGACAGAGAAACAAATTTGGATCACCCTTGTCAAGGTTAATTCATTGGAGGCCA 1520
DB 1402 AAGTGGATATGATGAAGCAAAAGGTTTACAAATGCTTAAAGCCAAAGCCCTTTGGAAGCCA 1461

QY 1521 TGTGCAAGGCTCGCCCACTTGGCACTTAAAT 1551
DB 1462 TGTGCAAGCAGCTCCCTCTGTGTTAAAT 1492

```

RESULT 9

AAC47389
ID AAC47389 standard; DNA; 1655 BP.

XX AAC47389;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 53631.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX PF1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 18-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131149.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149502.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.

PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	18-OCT-1999;	99US-0159638.	
PR	21-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 20.7%; Score 384.6; DB 21; Length 1655;			
Best Local Similarity 56.6%; Pred. No. 3.9e-80;			
Matches 849; Conservative 0; Mismatches 619; Indels 33; Gaps 6;			
QY	85	TGTCCTTTCTCGTTCCTCTCTCGCGTGAAGTTGTTTCCAAAGCAGAAAATTGAG	144
DB	77	TCTCCCTCTCTATTTCTTGTATATAGTTACAAATTTCTCTACTCCAAACGCAACGTTT	136
QY	145	AAACATACCAACAGGTCCTCTCTCTCCATAATAGGAAACCTTAACCTCTCGAAAC	204
DB	137	CAACCTTCTCCAGGACCAACCGTCAGACCTTTCGTGGCCATCTCCACCTCATGAAAC	196
QY	205	GCCAAATCCACGTTCTTCCAAACGATGTCGAAACAGTAGCGCAACGTTGCTCTCG	264
DB	197	GCGATCCACGCTCTCTCCAAACGCTACTCCAAACGCTACTCCAAACGCTCTCTCT	256
QY	265	GTTCCGTTACGCTCTGCGCTGTGCTCTCTCTCCAAACGATACCAAGAAATGCTTCA	324
DB	257	TTTCGGCTCCGAGCGCTGCTGTGTATACATCCCTTCTCTCGCCCAAGATCTCTCA	316
QY	325	---CAAAACAGAGTTGCTTGGCCAAACCGGCTACCTTCTCTCGGGAAATACATCTT	381
DB	317	CGGCAAAACGACATCTCTCTCCAGCGCGCTCCAACTCACCGCAAAATACGTCG	376
QY	382	CTACAAACACCAACGCTAGCTCTCTCCAGCGGAGCACTGGCGCAACCTCCGCG	441
DB	377	ATACAAACCAACACCGTTGGTACGCTCTTACGCGACCACTGGCGTAACCTCCGCG	436
QY	442	CATCACCGCTCTCTCTCCAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	501
DB	437	CATTGCTCCAAAGAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	496

QY	502	CGACGAGCGAAGCGTCTGATCGAGAGTTCGCTGCGCAAGAACTCGAACGAGAA--	559
DB	497	AGACGAGATCTTACGGATGCTCACACGCTCTCACGTTTACACAAACCTCCACGAAAG	556
QY	560	---GAGTTTCGGCGAGTGGAGATTAGTTCCGATGTTCAACGACTTAACTTACAAACAT	615
DB	557	CAACGATTTCTACTCATCGAGCTAGAACCCTCTTATCTGATCTTAACTTCAACAT	616
QY	616	AATGAGGATGATATCGGGAAAGAGGTTTTCGAGAGGAGAGTGAGATGAAGACGTTGA	675
DB	617	AGTAAGAATGGTTACAGGGAAGAGATATTCGCGCAGCAGC-----TCAACAAACA	670
QY	676	GGACGAGGAGGTTTCAGAGAGACTGTGACAGAAATGTTGGAACTCATGGGCTTGGCTAA	735
DB	671	AGAAGCAGAGCTATTTCAAGAAGCTAGTCTACGATATCGCCATGTATAGCGGCTTAATCA	730
QY	736	CAAGGGAGATCACTTTCCTTCAGGTGTTTCGATTTTCAGAAATGTTGGAGAAGCGCTT	795
DB	731	TTCCGCTGATTACTTACCCTACTTAACTATTCGAA---ACAATTCGAGAAAGAGT	787
QY	796	AAAGAGTATCAGTAAGAGGTACGATTCCTCTTGAATAGATCTTTCATGAGAACCGTGC	855
DB	788	TAAAGCTATAGCAATCAATGGATGATATTTTCAGCGTTTGTGTAGTAA---TGTAG	844
QY	856	CAGCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCC	915
DB	845	AAGAGATAAAGAAGGTAAACAATGTCATCAATCATTTGATCTCTCTACCAACAACACC	904
QY	916	TCAGTACTACTGACCAAAATCATCAAAAGGCTCTCTGCGCATGCTTTTGGTGGAAAC	975
DB	905	TGAGTATTAAGTACGCTGATCATCAAAAGGCTTATGATGTCGATGATGCTGCCGGAAC	964
QY	976	TGACTCATCACTGGGACTTTAGAGTGGTCAATTAATTAATTTGAATCACCCAGAGT	1035
DB	965	TGAGACCTCCGCTGTGACATTTAGAGTGGCGGATGCGAATCTACTGAGAAATCCAGAGT	1024
QY	1036	GTTGAAGAGGCAAGAGATGAAATGGACACTCAAGTGGGCAAGACCGCTTGTAAATGA	1095
DB	1025	ACTGGAGAAGGCGAGATCGAGATCGATGAGAAGATCGGAAAGATCGAATAATCGACGA	1084
QY	1096	GTCAGACCTTCCAAACCTTCCATATCTTAGGAAGATCATCTTGGAGACATTAAGTTGTA	1155
DB	1085	ATCAGACATCGCAGTGTCTTCTTACCTCCAAACGCTGTCTCGAAACATTCGGCTATT	1144
QY	1156	CCCCCGGCGCCAAATCTAATACCTCATGTGTCTTTCAGAGATATTAACAATGAAGATT	1215
DB	1145	TCGGTGGCGCGTTTCTAATCCCTCGATCACCGACGACGACATGAAATCGAGGTTA	1204
QY	1216	CAATATCCACGAGACACAATTTGATCATTAATGTTGGGCGATCGAGAGATCCTCA	1275
DB	1205	CGACGTCCGCGAGATACGATAGTGTGTTAACGCTTGGGCTATACATAGAGATCCAGA	1264
QY	1276	GTTGTGGAATGATGCCATGCTTTAAACCTGAGAGGTTTGTGTGGAAGGAG-----	1328
DB	1265	GATTTGGGAGGACGAGAGAGTTTATCCAGATCGATATATGACGATGCGGAAGCGA	1324
QY	1329	-----AGGAAAAAGTTGGTAGCATTTGGCATGGGAAGAGGGCTTGCACGAGAAAC	1383
DB	1325	TTACTATGTTTCAAGCTGATGCCGTTTGGGAATGGCGGAGAACTTGTCCGCGCCCG	1384
QY	1384	CATGGCTATGAAAGTGTGAGCTTTTACCTTGGGATTTGTGATTCATGTTTGTACTGAA	1443
DB	1385	ATTAGTTCAGAGGATTTGACATTTGGCGCTTGGATCGTTGATTCATGCTTTGAAATGGGA	1444
QY	1444	ACGAGTAAAGTAAAGGCTTGTATATGACAGAGAACTTGGATCACCTTGTCAAGGTT	1503
DB	1445	GAAATGAAAGGGAAGAGATGATATGCTGAGAGTACTGGTGGGTATGCGTAGAT	1504
QY	1504	AATTCATTTGAGGCGCATGTGCAAGCTCGCCACCTTGCACCTAAATTTGGAATTTAAT	1563
DB	1505	GGATCCTTTACGGGCAATGTGTAGGCTTAGGCCCTATATGTCAAAGCTTTTACTCTAAATA	1564

QY 1564 A 1564
Db 1565 A 1565

RESULT 10
AAC37476
ID AAC37476 standard; DNA; 1656 BP.
XX AC
XX AAC37476;
XX DT 17-OCT-2000 (first entry)
XX XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17517.
XX XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX XX
OS Arabidopsis thaliana.
XX XX
PN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136332.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR	16-AUG-1999;	99US-0149368.	
PR	17-AUG-1999;	99US-0149175.	
PR	18-AUG-1999;	99US-0149426.	
PR	20-AUG-1999;	99US-0149722.	
PR	20-AUG-1999;	99US-0149723.	
PR	20-AUG-1999;	99US-0149923.	
PR	23-AUG-1999;	99US-0149902.	
PR	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151065.	
PR	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 20.7%; Score 384.6; DB 21; Length 1656;			
Best Local Similarity 56.6%; Pred. No. 3.9e-80;			
Matches 849; Conservative 0; Mismatches 619; Indels 33; Gaps 6;			
QY	85	TGTCCTTTTCTCGGTTCTATTCTCGCGTGAAGTTTGTTCCTCAAGCAGAAAATTGAG 144	
DB	77	TCCTCCCTCTCTATTCTTGTATAAGTTACAAAATTTCTACTCCAAAACGCAACGTTT 136	
QY	145	AAACATACCAACGAGTCTCTCTCTTCCCATTAATAGGAACCTTAACTCTCTCGAACA 204	
DB	137	CAACCTTCTCTCAGGACCAACCGTACGACCCCTTCGTCGGCCATCTCCACCTCATGAACG 196	

QY	205	GCCAAATCCACCGTTTCTTCCAAACGAGTGTGAAACAGTACGGCAACGTGGTTTCCCTCTG 264
DB	197	GCCGATCCACCGTCTCTCTCCAAACGCTACTCTCAACCAATAACGCGCCAAATCTTTTCCCTCG 256
QY	265	GTTGGTTTCAAGTCTGGCGGTTGTCATCTCTCTCCAAACGAGTGTGAAACAGTGTGGTTTCCAC 324
DB	257	TTTCCGCTCCCGACGCGTGTGTCATTAACATCCCTTCTCCGCGCAAGAAATCTCTTCCAC 316
QY	325	---CAAACACGACGTTGCTTGGCCAAACCGGCTACCTTCTCTCTCGGAAATAATACATTT 381
DB	317	CGGCCAAAACGACATCGTTCTCTCCAGCGCGCGCTCCAACTCACCTCCAAATACGTGCG 376
QY	382	CTAACAAACACGACGTTAGGCTCTCTCTCCAGCGGAGCACTGGCGCAACCTCCGCGG 441
DB	377	ATACAAACCAACAACCGTTGGTACAGCTCTCTTACCGGCGACCACTGGCGTAACCTCCGCGG 436
QY	442	CATCACGCGCTGAGAGCTCTCTCCAGCGCGGTCCACTCTCTCTCCGGAATCCGAG 501
DB	437	CATTTGCTCCCAAGAGATTTCTCTCTCCAGCGCTCTCACTCTCTCACTCTCAACATCCGCA 496
QY	502	CGACGACGAAAGCTCTGATGACGAGGTTGGTGTGGCCAAAGAACTCGAACGAGAA-- 559
DB	497	AGACGAGATCTTACGGATGCTCACAGGCTCTCAGTTTACACACAAACCTCCACGAAAG 556
QY	560	---GAGTTTGGCGAGTGGAGATTAGTTTCAGATGTTTCAACGACTTAACTTACAAACAAT 615
DB	557	CAACGATTTCACTCAGCTAGAGTACCGCTCTTATCTCTGATCTTAACTTCAACAACAT 616
QY	616	AATGAGGATGATCGGGGAGAGGTTTACGGAGAGGAGTGTGAGATGAGAGAGTTGA 675
DB	617	AGTAAGATGTTTACGGGAAAGAGATTTACGGGACGACG-----TCAACACAAGA 670
QY	676	GGAACGAGGAGTTTACGAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGGCTAA 735
DB	671	AGNAGCAGACTATTCAAGAGTAGTCTACGATATCGCATGTATAGCGGCTTAATCA 730
QY	736	CAAGGAGATCACTTGGCTTCTCAGGTGTTTTCAGATGTTTTCAGATGTTGAGAGGCTT 795
DB	731	TTCCGCTGATTTACTTACCAATCTAAAACTATTTCGAA--ACAAATTCGAGAAAAGT 787
QY	796	AAAGAGTATCAGTAAGAGGTACGATTCATCTTGAATAAGATCTCTTCATGAGAACCGTGC 855
DB	788	TAAAGTATAGCAAAATCAATGATGATATTTTGCAGCGTTTGTGTAGTAA--TGTAG 844
QY	856	CAGCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCC 915
DB	845	AAGAGATAAAGAGGTAACACAAATGGTCAATCATTTGATCTCTCTACAAACAAACAACC 904
QY	916	TCAGTACTACACTGACCAAAATCATCAAGGCTTCTCTGGCCATGCTTTTGTGGTGAAC 975
DB	905	TGAGTATTACACTGACGTTTATCATCAAGGCTTATGATGTCGATGCTCGCCGGAAC 964
QY	976	TGACTCATCAACTGGGACTTTTAGAGTGTCTATTATCTAAATTTATTGAAATCACCCAGAGT 1035
DB	965	TGAGACCTCCGCTGTGACATTAGATGGGCGATGCGAATCTACTGAGAAATCCAGAGT 1024
QY	1036	GTTTGAAGAGGCAAGAGATGAAATTGGACACTCAAGTGGGACAAAGACCGCTTGTAAATGA 1095
DB	1025	ACTGGAGAAGCGAGATCGGAGATCGATGAGAAGATCGAAAAGATCGACTAATCGACGA 1084
QY	1096	GTGAGACCTTCCAAAACCTTCCATATCTTAGGAAGATCATCTCTTGGACACATTAGTGTGA 1155
DB	1085	ATCAGACATCGCAGTGTCTCTTACCTCCAAAACGCTCTCCGAAACATTCGCGCTATT 1144
QY	1156	CCCCCGGCCCAATTTCTAAATACCTCATGTGTCTTTCAGAAGATATTTACAAATTAAGAGATT 1215
DB	1145	TCGGTGGCGGCTTCTTAATCCCTCGATCACCGAGGACGACATGAATAATCGAGGTTA 1204
QY	1216	CAATATCCCAACGAGACAAATTTGATCATTAATGTTGGGCGATCGAGAGATCTCTCA 1275
DB	1205	CGACGCTCCGCGAGATACGATAGTGTGTTAAACGCTTGGGCTATACATAGAGATCCAGA 1264

PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.	Query Match	20.6%;	Score 383.8;	DB 21;	Length 1545;
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.	Best Local Similarity	57.2%;	Pred. No. 5.9e-80;		
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.	Matches 794;	Conservative	0;	Mismatches 557;	Indels 36;
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.					Gaps 4;
PR	22-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160980.					
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.					
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.					
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.					
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.					
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.					
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.					
PR	28-JUL-1999;	99US-0145921.	PR	26-OCT-1999;	99US-0161360.					
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.					
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.					
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.					
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.					
PR	04-AUG-1999;	99US-014704.	PR	28-OCT-1999;	99US-0161993.					
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.					
PR	05-AUG-1999;	99US-0147192.								
PR	05-AUG-1999;	99US-0147260.								
PR	06-AUG-1999;	99US-0147303.								
PR	06-AUG-1999;	99US-0147416.								
PR	09-AUG-1999;	99US-0147493.								
PR	09-AUG-1999;	99US-0147935.								
PR	10-AUG-1999;	99US-0148171.								
PR	11-AUG-1999;	99US-0148319.								
PR	12-AUG-1999;	99US-0148341.								
PR	13-AUG-1999;	99US-0148565.								
PR	13-AUG-1999;	99US-0148684.								
PR	16-AUG-1999;	99US-0149368.								
PR	17-AUG-1999;	99US-0149175.								
PR	18-AUG-1999;	99US-0149426.								
PR	20-AUG-1999;	99US-0149722.								
PR	20-AUG-1999;	99US-0149723.								
PR	20-AUG-1999;	99US-0149929.								
PR	23-AUG-1999;	99US-0149902.								
PR	23-AUG-1999;	99US-0149930.								
PR	25-AUG-1999;	99US-0150566.								
PR	26-AUG-1999;	99US-0150884.								
PR	27-AUG-1999;	99US-0151065.								
PR	27-AUG-1999;	99US-0151066.								
PR	27-AUG-1999;	99US-0151080.								
PR	30-AUG-1999;	99US-0151303.								
PR	31-AUG-1999;	99US-0151438.								
PR	01-SEP-1999;	99US-0151930.								
PR	07-SEP-1999;	99US-0152363.								
PR	10-SEP-1999;	99US-0153070.								
PR	13-SEP-1999;	99US-0153758.								
PR	15-SEP-1999;	99US-0154018.								
PR	16-SEP-1999;	99US-0154039.								
PR	20-SEP-1999;	99US-0154779.								
PR	22-SEP-1999;	99US-0155139.								
PR	23-SEP-1999;	99US-0155186.								
PR	24-SEP-1999;	99US-0155659.								
PR	28-SEP-1999;	99US-0156458.								
PR	29-SEP-1999;	99US-0156596.								
PR	04-OCT-1999;	99US-0157117.								
PR	05-OCT-1999;	99US-0157753.								
PR	06-OCT-1999;	99US-0157865.								
PR	07-OCT-1999;	99US-0158029.								
PR	08-OCT-1999;	99US-0158232.								
PR	12-OCT-1999;	99US-0158369.								
PR	13-OCT-1999;	99US-0159293.								
PR	13-OCT-1999;	99US-0159294.								
PR	13-OCT-1999;	99US-0159295.								
PR	14-OCT-1999;	99US-0159323.								
PR	14-OCT-1999;	99US-0159330.								
PR	14-OCT-1999;	99US-0159331.								
PR	14-OCT-1999;	99US-0159637.								
PR	14-OCT-1999;	99US-0159638.								
PR	18-OCT-1999;	99US-0159584.								
PR	21-OCT-1999;	99US-0160741.								
PR	21-OCT-1999;	99US-0160767.								

QY 945 GCCTTCTCTGGCCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGAGTGT 1004
Db 932 GAATCATACTTACCTCTGATAATTTGACGGGACAGATACGTCTACAGTAACACTAGAAATGG 991
QY 1005 CATTATCTAATTTATTGAATCATCCCAAGAGGTGTTGAAGAAGCAAGAGATGAATTTGGACA 1064
Db 992 CAATGTCAAAATCTGTTGAACCATCCAGAAATCTTAAGAAAGCGAGATGGAATTCGATG 1051
QY 1065 CTCAAGTGGGACAAAGCCGCTTGTAAATGAGTCAGACCTTCCAAAATTTCCATATCTTA 1124
Db 1052 AAAAGTGGTTTAGACCGATTAGTAGACGAATCGGACATTTGTAATCTCTCTTATCTCC 1111
QY 1125 GGAAGATCATCTTTGAGACATTTAGTTGTACCCCGCCGCCAATTTCTAATACCTCATG 1184
Db 1112 AAAGCATTTGTTGGAACACATACGTATGTACCCGCGAGTCCCACTACTACTACCTCAT 1171
QY 1185 TGTCTTCAGAAGATATTACAATTTGAAGGATTCAATATCCACGAGACACAATTTGTATCA 1244
Db 1172 TGTATCAGAAGATTGTAAGTTGGAGGCTACGATATCAAGTGGAAACATTTGATTTGA 1231
QY 1245 TTAATGTTGGGCGATCGAGAGATCTCTCAGTTGTGGAATGATGCCATGCTTTAAAC 1304
Db 1232 CCAACCATGGGCCATGATAGATCCAGAGTATGGGAAGATCTTGAGATATTCARAC 1291
QY 1305 CTGAGAGTTTGTATGTGAAGAGAGAGAGAGAAAGTTGGTAGCATTTGGCATGGGAAGAA 1364
Db 1292 CAGAAAGATTGGAAGAGAGAGAGAGCTGAGAAGCTAATCTCATTTGGGATGGGACGAA 1351
QY 1365 GGGCTTGGCCAGGAGAACCATGCTATGCAAGTGTCAAGTGTACGTTTGGGATTTGCA 1424
Db 1352 GAGCTTGTCTGGGCGGGCTAGCTCATCGGCTAATAACCAAGCTCTTTGGAAGTTTGG 1411
QY 1425 TTCAATGTTTGAAGTGAACAGAGTAAGTGAAGGAAAGCTTGTATGACAGAGAAACAAT 1484
Db 1412 TTCAATGTTTGAAGTGAAGAGAGTTGGTGAAGATTTTGGACATGACCGAGACAAAG 1471
QY 1485 GGAATCCTTGTCAAGGTTAATTTCCATTTGAGGCCATGTGCAAGGCTCGCCACTTGGCA 1544
Db 1472 GAGCCACATTTGCCAAAGCTATACCATTAAGAGCCATGTGCAAGCAGCTTCTATTGTTG 1531
QY 1545 CTAAAT 1551
Db 1532 ATAACT 1538

RESULT 13

ABZ13668
ID ABZ13668 standard; DNA; 1503 BP.

AC ABZ13668;

21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1473.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216555-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US26685.

XX 24-AUG-2000; 2000US-227866P.

XX 26-JAN-2001; 2001US-284647P.

XX 22-JUN-2001; 2001US-300111P.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX

PI Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses

PS Claim 144; SEQ ID NO 1473; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.

CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

XX Sequence 1503 BP; 429 A; 391 C; 322 G; 361 T; 0 other;

Query Match 20.6%; Score 383.2; DB 24; Length 1503;

Best Local Similarity 56.5%; Pred. No. 8.1e-80;

Matches 847; Conservative 0; Mismatches 618; Indels 33; Gaps 6;

QY 85 TGTCTTTCTGTTCTATTCCTCGGCGTGAAGTTTGTTCCTCAAGCAGAAATTTGAG 144

Db 18 TCTCCCTCTTCTATTTGTAATAAGTTACAAAATTTCTACTCCTCAAAACGCAACGTTT 77

QY 145 AAACATACACACAGTCTCTCTCTCCATATAGGAAACCTTTAACTCTCGAAACA 204

Db 78 CAACCTTCTCAGAGACACCGTCAAGACCTTGTGCGCCATCTCCACCTCATGAAACC 137

QY 205 GCCAATCCACCGTTTCTTCCAAACGATGTGAAACAGTACGGCAACGTGGTTTCCCTCTG 264

Db 138 GCGGATCCACCGTCTCTCTCCAAACGCTACTCTCAACCAATACGCGCCAACTTTTCCCTCG 197

QY 265 GTTGGTTACGTCTGGCGGTTGTCTCTCTCCACAGCATACCAAGATGTTCTAC 324

Db 198 TTTCCGCTCCCGCGCTGCTGCTATCAATCCCTTCTCTCGCCCAAGAAATCTCTCAC 257

QY 325 ---CAACACAGAGTTGCTTTGGCCACCGGCTACTTCTCTCGGAAATAATACATCTT 381

Db 258 CGGCCAAACAGCATCGTTCTCTCAGCGCGCGCTCCAACTCACCGCCAAATAGTCGC 317

QY 382 CTACAAACACACCGTAGGCTCTGTCTCCACGCGAGCATCTGGCGCAACCTCGCGCG 441

Db 318 ATACAAACACACACCGTTGGTACAGCTCTTACGCGACCACTGGCGTAACCTCGCGCG 377

QY 442 CATCACCGGCTTGGAGCTCTCTCCACGAGCGGTCCACTCTCTCGGATCCGGAG 501

Db 378 CATTTGCTCCCAAGAGATTTCTCTCTCACCGCTCTCATCACTTCCAAACATCCGCA 437

QY 502 CGACGAGACGAGCGCTGTGATGCGAGGTTGTGTGCGCCAGAAACTCGAACGAGAA-- 559

Db 438 AGACGAGATTTACGGATGCTCACAGGCTCTCACGTTTACACAAACCTCCACGAAAG 497

QY 560 ---CAGTTTGGCGAGTGGAGATTGATTTGATGTTTCAACGCTTAACTTACAAACAT 615

Db 498 CAACGATTTCACTCATATCGAGCTAGAACCCTCTTATCTGATCTTAACTTCAACAACAT 557

QY 616 AATGAGGATGATATCGGGAAGAGGTTTTCAGGAGGAGATGAGATCAAGACGTTGA 675

Db 558 AGTAAGAATGGTTACAGGGAAGAGATATTACGCGCACGACG-----TCAACAAAGA 611

QY 676 GGAACGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAATCTATGGGCTTGGCTAA 735

XX

[illegible]

DT 17-OCT-2000 (first entry)

ID AAC41374 standard: DNA: 1653 BP.

AAC41374:

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 31648.

KW Hybridisation assay; genetic mapping; gene expression control;

protein identification; signal transduction pathway; KW

KW metabolic pathway; promoter; termination sequence; ss.

```
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.5%; Score 381; DB 21; Length 1653;
Best Local Similarity 57.5%; Pred. No. 2.7e-79;
Matches 797; Conservative 0; Mismatches 560; Indels 30; Gaps 5;

QY 171 TTCCCATATAGGAACCTTAACCTCTCGAACAGCCAAATCCACCGTTCTTCTTCCACGCA 230
DB 132 TACCGGTGATTGGTCACTCCGTCTTCTCAAAACACCGATTTCGCACATTCCTCTCC 191
QY 231 TGTGAAACAGTACGGCAACG-----TGGTTTCCTCTGGTTTCGGTTCACTCTGGCGG 284
DB 192 TCTCTCAATCCCTAAACAATGTCGATCTTCTCCCTCCGACTCGTAACCGACTCGTTT 251
QY 285 TTGTCTATCTCTCTCCAAACAGCATACCAAGATGTTTCCACCAACACGAGTTGCTTCGG 344
DB 252 TCGTGAACCTGTCACACTCGATCGCGGAGGAATGTTTCCACCAAAACGAGTCGTA 311
QY 345 CCAACCGGTACTCTCTCTCGGAAAATACATCTTCTACAAACAACACCGTAGGCT 404
DB 312 CGAACAGACCAACTTCACTCTCGTAAACAGTTGCGTACGATTACACAACCATATCG 371
QY 405 CCTGCTCCACGGGAGCACTGGCGCAACCTCCGGCGATACCGCCCTGGAGCTCTCT 464
DB 372 CAGCTTCTACGGTGACCACTGGCGTAACCTCCGGCGATCGCTCCGTCGAGATATCT 431
QY 465 CCACGAGCGGTCCACTCTCTCCGGAATCCGGAGCGACGAGACGAGCGTCTGATCG 524
DB 432 CCAACCGGTCTCAATAGCTTTCTGTCTATTTCGTAAGACGAGATCCGACGACTGTGT 491
QY 525 AGAGGTGTGTGCTGCGCCAAAGRACTCGAACGAGGAAGATTTTCGCGAGTGGAGATTAGTT 584
```

Db 492 TTGGTCTTTTACGTAACCTTTT-----CACAGAGTTTGTGAAGTGGATATGAAT 542
Qy
585 CGATGTTTCAACGCTTAACCTTACAAACATAATAGGATGATATCGGGGAAGAGTTT 644
Db
543 CAATGTTATCTGACTTAACATTCACACACATTATAGAAATGGTGGCGGAACGTTACT 602
Qy 645 ACGGAGGAGAGATGAGATGAAGAACCTTGAGGAAGCGGAGGATTCAGAGACTGTGA 704
Db 603 ACGGAGACGGTGTG-----AGGATGATCCGAGGCTAAACGCTGTCGGCAGCTATAG 656
Qy 705 CAGAAATGTTGGAACCTCATGGCTTGCCTAACAGGAGATCACTTCCCTCAGT 764
Db 657 CGGATGGTGGCTGTGCTGGTCTGGAACGCTGTTGATTAATACCGGTTTGGCGT 716
Qy 765 GGTTCGAATTTTCAGAATGTGGAAGCGCTTAAAGAGTATCAGTAAGAGGTACGAATCCA 824
Db 717 TG-----GTTTCAGATTACGAACACGGGTAAAGAGTTAGCGGTAGGCTCGACGAT 770
Qy 825 TCTTGAATAGATCTTTCATGAGAACCGTGCAGCAATGACCGCCAGAAATTCATGATCG 884
Db 771 TCTTGAAGGATTTGGTTGATGAGAAACGAGACGCTAAGGAGAAAGGAAACACTATGATTG 830
Qy 885 ATCATCTCTCAACACTCAAGAGACCCAGCTCAGTACTACCTGACCTCAACCAATCATCAAG 944
Db 831 ATCATCTGCTTACTCTCAAGATCAACACCGGATTAATCTCAGTATGATCATCAAG 890
Qy 945 GCCTTGCTCGCCATGCTTTTGGTGAACCTGACTCATCACTCACTGGGACTTTAGAGTGT 1004
Db 891 GAAACATGCTTGCTTTGATCTAGCAGGACCGACACATCAGCGGTTACGTTAGATGG 950
Qy 1005 CATTATCTAATTTAATGAATCACCAGAGGTGTGAAGAGGCAAGAGATGAATGGACA 1064
Db 951 CATTGTCGAAGTGTGTTGAACCATTCGGAGGTATTGAACAAGGCGAGAGATGAATCGATA 1010
Qy 1065 CTCAGTGGGACAGACCGCTGTGTAATAGTACAGACCTTCCAAACTTCCATATCTTA 1124
Db 1011 GAAAGATCGGTTAGACAGGCTTAAGATGAATCAGATATCTCAACCTGCTTAATCTCC 1070
Qy 1125 GGAAGATCATCTTTGAGACACTTAGTGTGTACCCCGCGGCCCAATTTCTAATACCTCATG 1184
Db 1071 AAAACATTGTGCTGAAACGTTGCGCTTTATCTGCGGCTCCCATGCTTCTTCTCAAG 1130
Qy 1185 TGTCTTCAAGAGATATTACAAATGAAGATTCATATCCACGAGACACAAATTTGTATCA 1244
Db 1131 TTGCTCGAAGATTGTAAGTTGCAGGATATGATATCGCGGTGGCAGCATACTATTGA 1190
Qy 1245 TTAATGTTGGGATCGACAGAGATCTCAGTTGTGAATGATGATCCACATGCTTTAAAC 1304
Db 1191 CCAATGTGGGCTATACACAGATCTCTAGTATGGATGATCCAAATGAGCTTCAAGC 1250
Qy 1305 CTGAGAGTTTGTATGTGAAGGAGAGAGAAAGTTGGTAGCAATTTGGCATGGGAAGAA 1364
Db 1251 CAGAGAGTTTGAAGAGGAGAGAGCTCAGAGACTTATGCGGTTTGGGTAGGAAGAA 1310
Qy 1365 GGGCTTCCAGGAGAACCCATGGCTATGCAAGTGTGAGCTTACTTTGGATTTGATGA 1424
Db 1311 GGGGCTGTCTGGTCTGTGGACTGCTCACCGGCTTATAAACCTGACTCTTGGATCAATGA 1370
Qy 1425 TTCAATGTTTGTGATGGAACAGAGTAAGTGAAGGAAAGCTTGTATGACAGAGAACTAT 1484
Db 1371 TTCAGTGTGGAATGGGA----GAAAGATTGGAAGAGAGTGGATATGAGTGAAGGAAAG 1427
Qy 1485 GGATCAGCTTGTCAAGGTTAATTCATTGGAGGCGCATGTGCAAGGCTCGCCCACTGCCA 1544
Db 1428 GTGTTACATCCCTAAAGCCAGGCTTTAGAGGCCATGTGACAGAGCAGCTCCCTCTGTTG 1487
Qy 1545 CTAAGAT 1551
Db 1488 TTAAGAT 1494

RESULT 15
ABZ12256

ID ABZ12256 standard; DNA; 1503 BP.
XX
AC ABZ12256;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 61.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26685.
XX
PR 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-26467P.
PR 22-JUN-2001; 2001US-30011P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
XX WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
and producing plants with increased tolerance to these abiotic stresses
-
XX
PS Claim 144; SEQ ID NO 61; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
in the plant cell with an array or probes representative of the plant
cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
in methods of the invention.
CC
CC Note: The sequence data for this patent is not represented in the printed
specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 1503 BP; 397 A; 348 C; 374 G; 384 T; 0 other;
Query Match 20.3%; Score 377; DB 24; Length 1503;
Best Local Similarity 58.1%; Pred. No. 2.3e-78;
Matches 810; Conservative 0; Mismatches 550; Indels 33; Gaps 7;
Qy 171 TTCCCAATAATAGGAACCTTAACCTCTCGAAGCAAGCAATCCACCGTTTCTTCCAAACGCA 230
Db 113 TACCGGTGATCGTACCTCCGCTTCTCAAGCGGCACTCCACCGAGTTTCTATCCG 172
Qy 231 TGTGGAACAGTACGCAACG-----TGGTTTCCCTCTGGTTCCGTTACGTTCCGCGC 284
Db 173 TCTCTGAGTCCCTTAGCGGATGCTCCGATTAATCTCCCTGCTCTCGGCAACCGACTCGTTT 232
Qy 285 TTGTCACTCTCCCTCCACAGCATACCAAGATGCTTCCACCAACAGCGTTGCGCTGG 344
Db 233 TCGTTGTCTCTTCACTCACTCACTCGCGGAGATGCTTCAAGAGAGAGCTGCTACTCG 292
Qy 345 CCAACCGGCTACCTTCTCTCGGGAATAATACATCTTCTACAAACACACCGTAGGCT 404
Db 293 CGAACCGCTTCACTCCCTTGTCTTCAAAACACATCTCATATGGCTGCACACGGTTGTTA 352
Qy 405 CTGTGCTCCACGCGGAGCACTGCGGCAACCTCGCGCGCATCAACCGCCCTCGAGCTCTCT 464

Db 353 CAGCTTGTACGGGACCAATGGAGAAATCTCGCGGTATTTGGCGTGTGTGAGATATCT 412
Qy 465 CCACGACGCGGTCCACTCTCTTCGCGAATCGGAGCGACGAGACGAAGCGTCTCATGC 524
Db 413 CGGCTCACCGTCTCAATAGCTTTTCATCTATCCGCGAGAGAGATCCACCGGCTTAT-- 470
Qy 525 AGAGTTGGTGTGGCCCAAGAACTCGAACAGAGAAAGTTTGGCGAGTGGAGATTAGTT 584
Db 471 --AGCTGTCTCTCACGGAACTCTTCACT-----CGAATTTACAAAGGTTGAGATGAAAT 523
Qy 585 CGATGTTCAACGACTTAACTTTACAACAATAATAGGATGATATCGGGGAAGAGGTTTT 644
Db 524 CAATGTTTTCTAATCTTAACTTCAACAACATTAATAAGATTTGGCTGGAAATGTTACT 583
Qy 645 ACCGAGAGAGAGTGAATGAAGAACGTTGAGGAACGAGGAGTTTCAAGAGACTGTGA 704
Db 584 ACCGAGA-----TGGTGACAGAGATGACCCGAGGCTAAACGCGTGGGAGCTTATAG 637
Qy 705 CAGAAATGTTGGAATCTCATGGCTTGGCTTAAACAGGAGATCACTTGGCTTTCTCAGGT 764
Db 638 CAGAGGGAATGGGTGTTGTTTGGTGTGCTGGAACACTGCTGATTTATTTACCAGTCTGACTT 697
Qy 765 GGTTCGATTTTCAGAAATGTCGAGAACGCTTAAAGAGTATCACTAAGAGTACGATTCCA 824
Db 698 GGATC-----ACAGGTTCCGAGAAACGGATAAAGAGATCGCTAGTAGGCTTGTAGT 751
Qy 825 TCTTGAATAAGATCCTTTCATGAGAACCGTGCACGCAATG--ACCGCCAGAATTCATGA 881
Db 752 TCTTGCAGGATTTGGTGCATGAAGACGAGAAAGGAGGAGAAAGGACGACAAATGG 811
Qy 882 TCGATCATCTCCAACTCAAGACAGACCAGGCTCAGTACTACTACCTGACCAAAATCATCA 941
Db 812 TTGATCACTTGTCTGCTCAAGAAACCCAAACCCGAGTACTACCGGACAAACATTTATCA 871
Qy 942 AAGGCTTGTCTCGGCCATCTTTTGGTGGAACTGACTCATCAACTGGGACTTTAGAGT 1001
Db 872 AAGGAATCATGCTCTCTTTATCTAGCGGGGACTGATCATCCGAGTAACATTTGGAAT 931
Qy 1002 GGTCAATTATCTAATTTTAAATCACCAGAGGTGTTGAAGAAGGCAAGAGATGAATTTGG 1061
Db 932 GGACATTTGCAGCTTTGTTGAACCACTCCACAGATACTAAGCAAGGCGGAGAGATGAATCG 991
Qy 1062 ACCTCAAGTGGGACAAAGACCGCTTGTAAATCAGTCAAGACCTTCCAAAACCTTCCATATC 1121
Db 992 ATAATAAGTTGGCTCAACAGGCTTGTAGAGGAATCAGACCTCTCACATCTTCTTATC 1051
Qy 1122 TTAGGAGATCATCTTGAGACACTTAGGTTGTACCCCGCGCCCAATTTCTAATACCTC 1181
Db 1052 TTCAAAACATTTGTCCGAGTGTGGCTTTATACCCCTGCGAGTCCCTTGTGGTTCCTC 1111
Qy 1182 ATGTGTCTTCAGAAAGATATTACAATTTGAAGGATTTCAATATCCACGAGACACAAATTTGTA 1241
Db 1112 ACCTAGCATCAGAAGACTGTAAGTAGGGGTTTACCATATGCCACGTGGACGATGCTAT 1171
Qy 1242 TCATTAATGTTGGGGCATCGAGAGATCTCTCAGTTTGGAAATGATGACCATGCTTTTA 1301
Db 1172 TGACTAATGATGGGCAATACATAGGATCTTAAGATATGGGATGATGCCAACGAGCTTCA 1231
Qy 1302 AACCTGAGAGTTTGTATGTGAGAGGAGAGAGAAAGTTGGTAGCATTTGGCATGGAA 1361
Db 1232 AACCAGAGAGTTTCGAAAGAAAGGGGAGGCTCAGAAGCTATTGGGGTTTGGGCTAGGAA 1291
Qy 1362 GAAGGGCTTCCGAGGAGAACCCATGGCTATGCAAAAGTGTCAAGCTTTTACTTTGGGATTTG 1421
Db 1292 GAAGGGCTGTCTGGGTCCGGTCTGGCTCAACGGCTAGCGAGCTTGACTATCGGGCTTT 1351
Qy 1422 TGATTCATGTTTGTGCTGAAACGAGTAAAGTGAAGAAAGCTTTGATATGAC---AGAGA 1478
Db 1352 TGATTCATGTTTGTGCTGAAACGAGTAAAGTGAAGAAAGCTTTGATATGAC---AGAGA 1411
Qy 1479 ACAATTGGATCACTTGTCAAGGTTTAATTCATTGGAGGCCATGTGCAAGGCTCGCCAC 1538

Db 1412 GAGGAGGAGTCAATAATGCCAAAGCTATACCGTTGGTAGCCATGTGCAAGCACGCCCCG 1471
Qy 1539 TTGCCACTAAAAAT 1551
Db 1472 TTGTTGGTAAAGAT 1484
Search completed: October 21, 2003, 02:54:39
Job time : 509 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	393.8	21.2	1488	10	US-09-938-842A-1188	Sequence
2	386.6	20.8	1494	10	US-09-938-842A-23278	Sequence
3	383.8	20.6	1479	10	US-09-938-842A-2172	Sequence
4	383.2	20.6	1503	10	US-09-938-842A-1473	Sequence
5	377	20.3	1503	10	US-09-938-842A-61	Sequence
6	361.8	19.5	1509	10	US-09-938-842A-835	Sequence
7	350.6	18.9	1557	10	US-09-938-842A-1959	Sequence
8	330.4	17.8	1548	10	US-09-938-842A-1541	Sequence
9	202.8	10.9	722	9	US-09-770-149-144	Sequence 1
10	192.6	10.4	923	12	US-10-259-165-565	Sequence
11	189	10.2	975	12	US-10-259-165-433	Sequence
12	189	10.2	978	12	US-10-259-165-101	Sequence
13	182.2	9.8	447	15	US-10-174-693-221	Sequence
14	181.6	9.8	1545	12	US-10-356-153-55	Sequence
15	181.6	9.8	1545	14	US-10-142-231-55	Sequence
16	173.4	9.3	981	12	US-10-259-165-743	Sequence

Qy 276 GTCTGGCGGTGTCTCTCTCCAAACAGATACCAAGAAATGTTTCCACAAACACGACG 335
Db |||||
Qy 227 GCCTCGTTTACGTCTATCTCTCCACATCCATAGCCGAAGAATGTTTCCACAAACGACG 286
Db |||||
Qy 336 TTGCTTTGGCCAAACGGGTACTTCTCTCGGGAATAATCATCTTCTAACAACACCA 395
Db |||||
Qy 287 TCGTTCTTGGGAACCGTCTGATATCATGATGCGCAACACGTCGGCTATAACTTTACCA 346
Db |||||
Qy 396 CCGTAGGCTCTGCTCCACCGCGAGCATGGCGCAACCTCCGCGCGCATCACCGCCCTCG 455
Db |||||
Qy 347 ATATGATTGACATCTTACCGCGACCACTGGAGGAATCTCCGCGCATGCGCGCGCTCG 406
Db |||||
Qy 456 ACGTCTCTCCACGACGGGTCTCACTCTCTCCGGAATCCGAGCGAGACGAAAGC 515
Db |||||
Qy 407 AGATATTCTCTCATAGATCATGATACCTTTCTTCTATCCGTAAGACGAGATCCGAC 466
Db |||||
Qy 516 GTCTGATGACAGGTTGGTGTGCGCAAGAACTCGAAGGAGAAAGATTTGCGCGAGTGG 575
Db |||||
Qy 467 GGCTCATAACACATCTCTTAGAGACT-----CCTTGACCGGATTTGTTGAAGTGG 517
Db |||||
Qy 576 AGATTAGTTGATGTTCAACGACTTAATCTTACAAACATAATGAGGATGATATCGGGGA 635
Db |||||
Qy 518 AGTTGAATCGTTTATTAACCAACTTGGCATTCACACATCATCATGATGGTAGCCGAA 577
Db |||||
Qy 636 AGAGTTTACGAGAGAGAGTGAAGAAAGCTTTGAGGAAGCGAGGAGTTCAGAG 695
Db |||||
Qy 578 AACGATATTACGG-----TACCGGTACAGAAGCAATGATGAAGCCAAAGCTCGTAGGG 631
Db |||||
Qy 696 AGACTGTGACAGAAATGTTGAACTCATGGCTTGGCTTAACAGGGAGATCACTTGCCTT 755
Db |||||
Qy 632 AACTTATAGCGGAGATAATGCCGCGCGGTCTTGGAATTTGGCTGATATCTTCCGT 691
Db |||||
Qy 756 TCCTCAGGTGTTCCGATTTTCAGAAATGTGGAAGAGCGCTTAAAGAGTATCAGTAAGAGGT 815
Db |||||
Qy 692 CCAATATATGGGTC-----ACAACTTTGAGAACCAAGACCAAAATCTTGGGGAATCGAC 745
Db |||||
Qy 816 ACGATTCCATCTGAATGAAGTCTCTCATGAGAACCGTGCAGCAATGACCGCCAGAAAT 875
Db |||||
Qy 746 TCGATAGAGTCTCTCAAAACTGTTGATGAGAAACGCTGCAGAGAAAGAAAGGTCAAA 805
Db |||||
Qy 876 CCATGATGATCATCTCTCAAACTGCAAGAGACCCAGCTCAGTACTACATCAACCAAA 935
Db |||||
Qy 806 CTTTGATCGATCATTTGCTTCTTCCAGAAACCGAACCCGAGTACTTACTACTGTC 865
Db |||||
Qy 936 TCATCAAGAGCGCTGCTCTGGCCATGCTTTTGTGGAATGACTCATCACTGGGACTT 995
Db |||||
Qy 866 TCATCAAGAGATCATCTGCTTGGTCTTTCGCGGAGCAGATACATCTTCAGTAGCT 925
Db |||||
Qy 996 TAGAGTGTCTAATCTAATTTAATGAATCAACGAGAGGTGTTGAAGAGGCAAGAGATG 1055
Db |||||
Qy 926 TGGAAATGGCAATGTCAAAATTTGTGAACCATCCAGAGATACTTGAAGAAAGCGAGCGG 985
Db |||||
Qy 1056 AATTGGACACTCAAGTGGGCAAGACCGCTTGTAAATGAGTCAGACCTTCCAAACATTC 1115
Db |||||
Qy 986 AGATCGATGATAAATCGGTTGACACCGGTTAGTTGAAGAAATCAGATATTGAATCTCC 1045
Db |||||
Qy 1116 CATATCTTAGAAGATCATCTTGAGACACTTAGGTTGTACCCCGCCGCCCAATTTCTAA 1175
Db |||||
Qy 1046 ATTATCTCCAAACATTTGTGTGAGAAACATTAAGTTGTATCTCGGTTCCACTACTAC 1105
Db |||||
Qy 1176 TACTCATGTCTTTCAGAGATATTAAATGAAGGATTCATATCCAGAGACACAA 1235
Db |||||
Qy 1106 TCCCTCATTTCTCATCGGATGAATGAAAGTGGCGGCTAGCATATGCGAGTCCGACGT 1165
Db |||||
Qy 1236 TTGTGATCATTAATGTTGGGCGATGACAGAGATCTCTCAGTTGTGGAATATGCCACAT 1295
Db |||||
Qy 1166 TGTATTAAACACGTATGGGCGATGCTAGAGATCCAGGTTTATGGAAGAGCCAGAGA 1225
Db |||||
Qy 1296 GCTTTAAACCTGAGAGGTTTGTATGGAAGAGAGAGGAAAGATTTGTTAGCATTTGGCA 1355
Db |||||
Qy 1226 GGTTCAGCCGAGAGGTTTCGAGAAAGAGAGGAGGCTCTGAAGACTTAATGCGGTTGGGA 1285
Db |||||
Qy 1356 TGGGAAGAGGGCTTGGCCAGAGAAACCATGGCTATGCAAAAGTGTACAGTTTACTTTGG 1415
Db |||||

Db 1286 TGGACGACGAGCTTGTCTCGAGCTGAGCTTTGGGAAGCGGTTAGTGAGCCCTTGCTCTTG 1345
Qy |||||
Qy 1416 GATTGTTGATTCAATGTTTGTACTGGAACGAGTAAGTGAAGAAAGCTTTGATATGACAG 1475
Db |||||
Qy 1346 GGTGCTTGAATCAGCTTTTCGAGTGGGAGAGAGTTGGTGCAAACTTGTGGACATGCTG 1405
Db |||||
Qy 1476 AGAACAAATTTGGATCACTCTTCAAGGTTAAATTTCCATTGAGGCGCATGTCAGAGGCTCGCC 1535
Db |||||
Qy 1406 AAGCGAGAGGATCACTATGCTAAAGCTACTCCGTTGGGAGCTATGTGCAAGGCAAGTG 1465
Qy |||||
Qy 1536 CACTTGGCACTAAAA 1550
Db |||||
Qy 1466 CCAATTGTTGGTAAAA 1480
Db |||||

RESULT 2

US-09-938-842A-2278
; Sequence 2278, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2278
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2278

Query Match 20.8%; Score 386.6; DB 10; Length 1494;
Best Local Similarity 57.0%; Pred. No. 2.8e-93;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

Qy 90 TTTTCTCTGTTCTTATTCCTCGGCGTGAAGTTTGTGTTTCCAAAGCAGAAATTCAGAAACA 149
Db 29 TTCTCTCTGTTGTTCTCTCTCTCAATTTACTTAAATGGAAAACTCAAGCGAAAGCCAAATC 88
Qy 150 TACCACGAGTCTCTCC---TCCTCTTCCCATATAGGAAACCTTAACTCTCGAAGCAGC 206
Db 89 TACTCTCGAGTCGGGATGGTCATTACCGGTGATTGTCATCTCCGCTTCTCAAAACCAC 148
Qy 207 CAATCCACCGTTTCTTCCAAACGATGCGAAACAGTACGGCAACG-----TGGTTTCCC 260
Db 149 CGATTTCATGCAATTCCTCTCTCTCAATCCCTAAACATGTCTCCGATCTTCTCCC 208
Qy 261 TCTGTTTCGGTTTCAGCTGCGGCTGTCATCTCTCTCCATCCCAAGCAGCATACCAAGATGCT 320
Db 209 TCCGACTCGGTACCCAGCTGTTTCTGTAACCTGTCACACTCGATCGCGCGAAGATGTT 268
Qy 321 TCACCAAAACACGAGTTCCTTGGCCAAACCGGCTACTCTCTCTCTCGGGAATAATCATCT 380
Db 269 TCACCAAAACGAGCTGCTACTGCGCAACAGACCAAACTTCATCTCTCGTAAACAGTTG 328
Qy 381 TCTTAAACAAACCAACCGTAGGCTCTCTGTCCTCCACGCGGAGCACTGCGGCAACCTCGGCC 440
Db 329 CGTACGATTACAAACCATGATCGAGCTTCTCTACGCTGACCACTGCGGTAACCTCGGCC 388
Qy 441 GCATCAGCGGCTGGAGGCTCTCTCCACGAGCGGCTCCACTCTCTCTCGGGAATCCGGA 500
Db |||||

389 GCATCGGCTCGGTCGAGATATCTCCAACTCACCGTCTCAATAGCTTTTGTCTATTCGTA 448
501 GCGACGAGCAAGCGCTGTGATCGAGGTTGGTCTGGCCCAAGAACTCGAACGAGGAAG 560
449 AAGACGAGATCCGACGACTGTGTTTCGTTCTTCGGAAC-----TTTTCACAAG 499
561 AGTTTCGGCGAGTGGAGATTAGTTCGATTTCAACGACTTTCAACGACTTAACCAACATATGCA 620
500 AGTTTGTGAAGTGGATGAATCAATGTTATCTGACTTAACATTTCAACCAATTTTAA 559
621 GGATGATATCGGGGAAGAGGTTTACGGAGAGGAGTGAAGATGAAGAAAGCTTTGAGGAAG 680
560 GAATGCTGCGCGAARACGTTACTACGGAGACGTTG-----AGGATGATCCGGAGG 613
681 CGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGCTTGGCTTAACAAGG 740
614 CTAACAGTGTCCGGCAGCTTATAGCGGATGTGGTGTGCTGTGCTGGAACGCTG 673
741 GAGATCACTTGCCTTTCCTCAGGTGTTGATTTTCAGAAATGTCGAGAGCGCTTAAAGA 800
674 TTGATTACTTACCGGTTTTCGGTTGG-----TTTCAGATTACGAGACCGGTTAAGA 727
801 GTATCACTAAGAGTACGATTCCATCTTGAATAAGATCCTTCATGAGAACCGTGCAGCA 860
728 AGTTGGCGGTAGGTCGACGAGTCTTGCAGAGATTGGTGTGATGAGAAACGAGAAGCTA 787
861 ATGACCGCCAGAAATCCATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCTTCAGT 920
788 AGGAGAAAGGAACACATATGATCGATCATCTTACTCTGCAAGAAATCCCAACCGGATT 847
921 ACTACACTACCAAAATCATCAAGGCTTGTCTGTGGCCATGCTTTTGTGGTGAACGTACT 980
848 ACTTACCGATCGATCATTAAGAGAAACATGCTCGCTTTGTATCTAGCAGGACCGACA 907
981 CATCAACTGGGACTTTAGAGTGTCTATCTAATTTATTTGAATCAACCCAGAGGTGTCA 1040
908 CATCAGCGTTAGTTAGAAATGGGCAATGTCGACGTTGTAACCATCCGATGTATTGA 967
1041 AGAAGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTTGTTAAATGATGAC 1100
968 ACAAGCGAGAGATGAATTCGATAGAAAGATAGTTTGTAGCAGGCTTATGGATGAATCAG 1027
1101 ACCTTCCAAATCTCCATATCTTAGNAGATCATCTTGGACACTTAGTTGTACCCCC 1160
1028 ATATCTCAACCTGCGCTTATCTCCAAACATTTGTCTGAAACGTTGCGCCCTTATCTCG 1087
1161 CGGCCCCCAATTTCAATACCTCATGTCTTTCAGAAGATATTACAATTTGAAGGATTTCAATA 1220
1088 CGGCTCCCATGCTTCTCTCAGCTTGCCTCGGAGATTTGTAAGTTGCGAGATACGATA 1147
1221 TCCCGAGACACAATTTGATCATTAATGTTGGGGCATTCGAGAGATTCCTCAGTTGT 1280
1148 TGGCGGTGCGACGATCTATTGACCAATGTTGGGCTATACACAGAGATCTCTCAGCTAT 1207
1281 GGAATCATCCATCTTAAACCTGAGAGTTTGTATGAGAGAGAGGAGGAGAAAGT 1340
1208 GGGATGATCCAAATGAGCTTCAAGCCAGAGAGGTTTGAGAAAGAGAGAGCTTCGAGAGC 1267
1341 TGGTACATTTGGCATGGGAAGAGGCTTTCGCCAGGAGAACCCATGGCTTATGCAAGTG 1400
1268 TAATGCGGTTTGGTTAGGAGAGAGGCGGTGCTCGTCTGAGCTGGCTCATCGGCTTA 1327
1401 TCAGCTTTATTTGGGATTTGATTTCAATGTTTTCAGTTTTCAGTTGGAACGAGTAAAGTGA 1460
1328 TAAACCTGACTCTTGGATCATTTGATTTGAGTTTGAATGGGA---GAAGATTTGGAGAAG 1384
1461 AGCTTATATGACAGAGAACAAATTTGATCACCTTGTCAAGTTTAAATTCATTGGAGGCCA 1520
1385 AGTGTGATATGAGTGAAGGCAAGGTTTACAAATGCTTAAAGCCAAAGCCCTTTGGAAGCCA 1444
1521 TGTGCAAGGCTCGCCCACTTGGCACTTAAAT 1551
1445 TGTGACAGACACGCTCCTCTGTGTTAAAT 1475

RESULT 3

US-09-938-842A-2172
; Sequence 2172, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2172
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2172

Query Match 20.6%; Score 383.8; DB 10; Length 1479;

Best Local Similarity 57.2%; Pred. No. 1.6e-92; Mismatches 557; Indels 36; Gaps 4;

Matches 794; Conservative 0; Mismatches 557; Indels 36; Gaps 4;

QY 171 TTCCCATATATAGGAAACCTTAACCTCTCGAAGACAGCAATCCACCGTTTCTTCCCAACGCA 230
DB 116 TCGCGGTGATCGACACCTCCATCTCTCAAGCCACCGCTTCAACGTACTTCTTCC 175
QY 231 TGTGAAACAGTACGGCAA-----CGTGGTTTCCCTCTGGTTCCGTTACGCTCTGGCCG 284
DB 176 TTTCAAATCCATCGAAATGCTCGGTCTTCCATCTCCGACTCGGAAATCGCCTCGTT 235
QY 285 TTGTCTATCTCTTCCAAACAGATACCAAGATGTTTCCAAACACAGAGCTTGCCTTGG 344
DB 236 ATGTCATCTCTTACGTTCCATCGCTGAAGATGTTTCAAAAGAACAGATGCTGTTCTCG 295
QY 345 CCAACCGGCTACCTTCTCTCGGAAATATACATCTTCAACAAACACACCGTAGGCT 404
DB 296 CGAACCGCCCAAGTTCAACATAAGTAAGCACTCGGCTACAAACGCCACCTTACTTCT 355
QY 405 CTGTCTCCACGCGGAGACCTGCGCAACCTCCCGCGCATCACCGCCCTGGAGCGTCTCT 464
DB 356 CGGCATCTTACGGCGATCATTTGGAGAAACCTCCCGCGCATAGCCCGCTCGAGATATTCT 415
QY 465 CCACGACGCGGTCCACTCTTCTCCGGAATCCGAGGACGAGAGCGTCTGTATGC 524
DB 416 CCATCTATAGACTCAATTTGGTTTCTGTATATTCGTAAGACGAGATCCGACGCTCATTT 475
QY 525 AGAGTTTGTGTGTGCCAAGAACTCGAACGAGGAGAGTTTTCGCGAGTGGAGATTAGTT 584
DB 476 CACATCTCTCTGTTGATT-----CCTTACACGGATTGTGAGAGTAGAGATGAAAA 526
QY 585 CGATGTTCAACGACTTAACCTTCAACAAACATAATGAGGATGATATCGGGGAAGAGTTT 644
DB 527 CATTTGTTAAACAAATTTGGCATCCCAACCAACCATCAGAAATGTTGGCCGGAAGAGATA 586
QY 645 ACGGAGAGAGAGTGAAGATGAAGACGTTTGAAGAGCGGAGGAGTTTCAGAGACCTGTGA 704
DB 587 TCGGTGAGG-----ACAACGATGACCTAACTCGTGAAGAACCTTTGTGT 631
QY 705 CAGAAATGTTGGAATCATATGGGCTTGGCTTAACAGGAGATCACCTTGGCTTTCTCAGGT 764
DB 632 CGGAGCGGTGACACGCGCGGTGAGGAAACCCCATTTGATTATCTTTCCATTTTACGTT 691

QY 765 GGTTCGATTTTCAGATGTCGAGAGCGCTTAAGAGATATCAGTAAGAGGTACGATTTCCA 824
Db |||||
QY 692 GGGTCTC-----GAGTTATGAAGAAAGAAATTTGGGAAATAGGTTTGTACGT 745
Db |||||
QY 825 TCTTGAATAAGATCCTTCATGAGAACCGTGCAGCAATGACCGCAGATTTCCATGATCG 884
Db |||||
QY 746 TTTTGCAGAAATTAGTCAGCAAAACGTCGGGAGAGGAAAGGTTGAACTATGATCG 805
Db |||||
QY 885 ATCATCTCTCAAACTCAAGAGAGCCAGCTCTAGTACTACATGACCAAAATCATCAAG 944
Db |||||
QY 806 ATCACTTCTGCTCTCAAGACATTTCAACCTGATTACTATACGGATGTCATCATCAAG 865
Db |||||
QY 945 GCCITGCTCGCCATGCTTTTGGTGAATGATCTCATCACTACCTGGGACTTTAGAGTGT 1004
Db |||||
QY 866 GAATCATACTTACCTGATATTTGAGGAGCAGATACGTCATGATACACTAGAAATGGG 925
Db |||||
QY 1005 CATATCTAAATTTATGATCAACCCAGAGGTGTTGAAGAGCGCAAGAGATGAATGGACA 1064
Db |||||
QY 926 CAATGTCAAATCTGTTGAACCATCCAGAAATACTTTAAGAAAGCGGATGGAATCGATG 985
Db |||||
QY 1065 CTCAGTGGGCAAGACCGCTTGTAAATGATGAGACCTTCCAAAATTTCCATATCTTA 1124
Db |||||
QY 986 AAAAAGTCGGTTAGACCGATTAGTAGACGAATCGACATTTGTAATCTCTCTTATCTCC 1045
Db |||||
QY 1125 GGAAGATCATCTTTGAGACACTTAGTGTACCCCGGCCCAATTTCTAATACCTCATG 1184
Db |||||
QY 1046 AAGCAITGTATGTTGAACACCTACGATGTATACCGGCGAGTCCCACTACTACTCTCAT 1105
Db |||||
QY 1185 TGTCTTCAGAAATATTAACAATTTGAAGGATTTCAATATATCCACGAGACACAATTTGTATCA 1244
Db |||||
QY 1106 TGTCAATCAGAAATTTGTAAGTTGGAGCTAGGATATACCAAGTTGGAAATGTTATGA 1165
Db |||||
QY 1245 TTAATGTTGGGCGATGAGAGAGATCTCTAGTTGTGGAATGATGCCACATGCTTTAAAC 1304
Db |||||
QY 1166 CCAACGATGCGCCATGATGAGATCCAGAGGTATGGGAAGATCTTGAGATATTTCAAA 1225
Db |||||
QY 1305 CTGAGAGTTTGTATGTCGAAGAGAGGAGAAAGTTGGTAGCATTTGGCATGGGAAGAA 1364
Db |||||
QY 1226 CAGAAAGATTTGAAAAGAGAGAGGCTGAGAGCTAATCTCATTTGGGATGGAGCAA 1285
Db |||||
QY 1365 GGGCTTGCCAGGAGAACCCATGGCTATGCAAGTGTGAGCTTTACTTTGGGATTTGTA 1424
Db |||||
QY 1286 GAGCTTGTCTGGAGCGGCTAGCTCATCGCTTAATAAACAGGCTCTTGGAAATTTGG 1345
Db |||||
QY 1425 TTCAATGTTTGTGCTGGAAACGAGTAAGTAGGAGGAAAGCTTGATATGACAGAGAACTT 1484
Db |||||
QY 1346 TTCAATGTTTGTGTTGGAAAGAGTTGTTGAGGATTTTGTGGACATGACCGAAGACAAAG 1405
Db |||||
QY 1485 GGATCACTTGTCAAGGTTAATTCATTTGGAGGCGCATGTCGAAGGCTCGCCCATTTGCCA 1544
Db |||||
QY 1406 GAGCCATATGCCCAAGCTATACCATTAAGAGCCATGTGCAAGACGCTTCTTATTTGTTG 1465
Db |||||
QY 1545 CTAAAT 1551
Db |||||
QY 1466 ATAACT 1472
Db |||||

RESULT 4

US-09-938-842A-1473
; Sequence 1473, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1473
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1473

Query Match 20.6%; Score 383.2; DB 10; Length 1503;
Best Local Similarity 56.5%; Pred. No. 2.3e-92;
Matches 847; Conservative 0; Mismatches 618; Indels 33; Gaps 6;

QY 85 TGTCTCTTTCTCTGTTTCTATTCTCTCGGCGTGAAGTTTGTTCCTCAAGCAGAAAAATTGAG 144
Db 18 TCTCCCTCTCTCTATTCTTGTAAATAGTTTACAAATTTCTCTACTCCAAAACGCAAGTTT 77
QY 145 AATCATACACACAGGTCCTCTCTCTTCCATATAGAAACCTTAACCTCTCGAACA 204
Db 78 CAACCTTCTCTCAGGACCAACCGTCAGAGCCCTTCGTCGGCCATCTCCACCTCATGAACC 137
QY 205 GCCAATCCACCGTTCTTCCACGCACTGTGAAACAGTAGTACGCAACGTTGTTTCCCTCTG 264
Db 138 GCGATCCACCGTCTCTTCAACGCTACTCCAAACCAATAGGCCCAATCTTTTCCCTCCG 197
QY 265 GTTCGGTTCACTGCTGCGGCTTGTCTCTCTTCCACAGCATATCAAGAATGCTTTCAC 324
Db 198 TTTTGGCTCCGAGCGGTCTGTGTCTATAATATCCCTTCTCTCGCCCAAGAATCTTTCAC 257
QY 325 ---CAAAACAGAGTTCCTTGGCCAAACGGGTACTCTCTCTCGGGAATAATACATCTT 381
Db 258 CGGCCAAAACGATCTCTCTCTCCAGCGCGCGCTCCAACTCACCGCCAAATACGTCGC 317
QY 382 CTACAAACACACACCGTAGGCTCTCTCCACGCGGAGCACTGGCGCAACCTCCGCCG 441
Db 318 ATACACACACAAACGTTGTTGATACGCTCTTACGGGGACCACTGGGTAACCTCCGCCG 377
QY 442 CATCACCGCTCGAGCTCTCTCTCCACGAGCGCTCTCTCTCTTCGGGAATTCGGAG 501
Db 378 CATTTGCTCCAAAGATTTCTCTCTCTCACGCTCTCATCACTTCCACACATCCGCA 437
QY 502 CGAGGAGACGAGCTCTGATGAGAGGTTGTTGTCGGCCAGAACTCGAACAGGAA-- 559
Db 438 AGACGAGATTTACGGATGCTCTACGCGCTCTCACGCTTACACACAACTCCCAACGAAAG 497
QY 560 ---GAGTTTGGCGAGTGGAGATTAGTTTCGATGTTCAACGACTTAACCTTACAAACAT 615
Db 498 CAACGATTTCACTCATCGAGTAGAACCGCTCTTATCTGATCTAACATTCACACAT 557
QY 616 AATGAGATGATATCGGGGAAGAGGTTTTCAGGAGGAGAGTGAGATGAAGAACGTTGA 675
Db 558 AGTAAGAAATGGTTACAGGGAAGAGATATTACGGCGACGACG-----TCAACAAACAAAGA 611
QY 676 GGAAGCGAGGAGTTTCAAGAGACTGTGACAGAAATGTTGGAATCATGCGGCTTGGTAA 735
Db 612 AGAAGCAGAGCTATTCAAGAAAGCTAGTCTACGATATGCGCATATAGCGCGCTAATCA 671
QY 736 CAAGGGAGATCACTTGCCTTTCTCAGGTGTTTTCGATTTTCAGAAATGTCGAGAACGCTT 795
Db 672 TTCGCTGATATTACTTACCAATATAAATATTTCGGAA---ACAAATTCAGAAAGAGT 728
QY 796 AAAGAGTATCAGTAAGAGGTAAGATTCATCTTTGAATTAAGATCTTTCATGAGAACCGTGC 855
Db 729 TAAAGCTATAGGCAAAATCAATGATGATATTTTGGACGCTTGTGTAGATGAAT---GTAG 785
QY 856 CAGCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACTGCAAGACCCAGCC 915
Db 786 AAGAGATAAAGAGGTAACAAATATGTTCAATCTCTCTCAACCAACAAACAAACC 845
QY 916 TCAGTACTACACTGACCAATCATCAAGGCTTGTCTGCGCCATGCTTTTGTGTGAAC 975
Db |||||

QY 1242 TCATTAAATGGTTGGGCGATCGAGAGATCTCTAGTTGTGGATGATGCCACATGCTTTA 1301
Db 1172 TGACATAATGATGGGCAATACATAGGATCTTAAGATATGGATGATCCAAACGAGCTTCA 1231
QY 1302 AACCTGAGAGTTGATGTGGAAGGAGAGAGAAAGTTGGTAGCATTTGGCATGGGA 1361
Db 1232 AACAGAGAGTTGCGAAGAAAGGGGAGGCTCAGAAAGCTATTGGGGTTTGGGCTAGGA 1291
QY 1362 GAAGGGCTTGGCCAGAGAAACCCATGCTATGCAAAAGTGTCAAGTTTACCTTTGGGATGT 1421
Db 1292 GAAGGGCTGTCCCTGGGTCCGGTCTGCTCAACGGCTAGCGAGCTTGAATCGGTCTT 1351
QY 1422 TGATTAATGTTTGAATGCAAGAAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1478
Db 1352 TGATTAATGTTTGAATGCAAGAAAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1411
QY 1479 ACAATTCGATCACTTGTCAAGGTTAATTCATTTGAGGAGGAGTCAAGGCTCGGCCAC 1538
Db 1412 GAGGAGAGTCAATGCTCAAGGTTAATTCATTTGAGGAGGAGTCAAGGCTCGGCCAC 1471
QY 1539 TTGCCACTAAAT 1551
Db 1472 TTGTTGTAAGAT 1484

RESULT 6
US-09-938-842A-835
; Sequence 835, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 835
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-835

Query Match 19.5%; Score 361.8; DB 10; Length 1509;
Best Local Similarity 57.6%; Pred. No. 1.4e-86;
Matches 755; Conservative 0; Mismatches 532; Indels 24; Gaps 5;

QY 245 GCGAACGTTGTTCCCTCTGTTCCGTTACGTTACGTTGCGGCTGTTGTCATCTCTCCAA-- 302
Db 211 GCGGCGGTGATGTCCTCCGACTTGGATCCGCTCTAGTCTACGTGGTGTGTCACATAAG 270
QY 303 -CAGATACCAAGATGTTCCACCAACACGAGCTTGGCCACCGGCTACCTTCT 361
Db 271 GTTGGCGAGAGGATGCTTTGGCAAGAACGACGCTGTTCTAGTAAACCGGCCACAGGTG 330
QY 362 CTCTCGGAAAATACATCTTACAAACACACCGGCTGCTCTGCTCCCGCGGAG 421
Db 331 ATCATCGAACAACATGTTGGTTATTAACAAATACATGATCGCTGACCTTACGGTAT 390
QY 422 CACTGGCGCAACCTCCGCGGATCACCGCCCTGGAGGCTCTCTCCAGCGAGCGTCCAC 481
Db 391 CACTGGCGTAACCTCCGCGCTCTGCACTATCGAGATCTTTTCAACACACCGGCTTAAC 450
QY 482 TCCTTCTCCGGAATCCGAGCGAGAGCAAGCGTCTGATGCGAGAGTTGGTCTGCC 541

RESULT 7

Db 451 TCGCTTTCTTTATGTCGTAACGAGAGTACGTGCGCTTAT----AAGCGCTCTCTCG 506
QY 542 AAGAACTCGAACGAGAGAGTTTCGCGAGTGGAGATTAGTTTCGATGTTCAACGACTTA 601
Db 507 TCTCGAGGAACCAAAAACCTGT-----TGTTGAATGAAACCAATGCTTATGGACTTG 561
QY 602 ACTTACAACAATAGGATGATATCGGGGAAGAGGTTTTCGAGAGAGAGAGTAG 661
Db 562 ACATTTCAACAATAGGATGATGATCGGAAACGATACCTACGTTGAAGAAACA--- 618
QY 662 ATGAAGAAGCTTGAGAAAGGAGGAGTTTCAGAGAGAGTGTGACAGAAATGTTGGNACTC 721
Db 619 ---ACCGATGAGGAAGCAAAAGCGTGTCCGTAAGTTGGTAGCAGATGTTGGGGCTAAC 675
QY 722 ATGGGCTTGGCTTAAACAAGGGAGATCACTTTCCTCAGCTGGTTCGATTTTCAGAAAT 781
Db 676 ACGAGTTGAGGCAACGCGGTGACTACGTTCCCATCTCAGACTG-----TTTCAAGT 729
QY 782 GTGGAGAGCGCTTAAAGAGTATCAGTAAGAGGTAGATTCCATCTTGAATAGATCTCTT 841
Db 730 TATGAAAATAGGGTTAAAAAGTTTAGGAGAGGAGACAGATAAGTTTTTACAGGTCTCAT 789
QY 842 CATGAGAACCGTCCGAGCAATGACCGCAGAAATTCATGATCGATCATCTCTCAACATG 901
Db 790 GACGATAAACGTTGGGCAACAAGAACCGGTACTACAATGATCGATCATTTGCTTGTCTC 849
QY 902 CAAGAGACCCAGCTCAGTACTACACTGACCAAAATCATCAAAAGCGCTTCTCTGGCCATG 961
Db 850 CAATAATCTGATTTAGTATTACTGATCAATCATCAAAAGGATCATATCATGATAATG 909
QY 962 CTTTTTGGTGGAACTGACTCATCAACTGGGACTTTTAGAGTGTCTATATCTAATTTATG 1021
Db 910 GTAATAGCAGGAGCTAAACATCAGCAGTCACTCTAGAAATGGGCACTTTTCCAATTTGCT 969
QY 1022 AATCAACCCAGAGGTGTTGAAGAGCAAGAGATGAATTCGACACTCAAGTGGGCAAGAC 1081
Db 970 AACCATCTCTGAGTAATATCTAAAGCGAGAGAGAAATCGATAACCGGGTGGTTTAGAC 1029
QY 1082 CGCTTTTAAATGAGTACAGACCTTCCAAAACTTTCCATATCTTAGGAAGATCATCTTGG 1141
Db 1030 CGCTTTATCGAAGAGCAGATCTTAGCAGAGTACCATACTAAAGAAACATTTGCTTGG 1089
QY 1142 ACATTTAGTGTGTAACCCCGGCGCCCAATTTAATACCTCATGTGTCTTCAGAAATATT 1201
Db 1090 ACCCTTAGGCTACACCCAGCTACACCAATGCTAGTCCCAACATGGCATCAGAAATTC 1149
QY 1202 ACAATTTGAAGATTCAATATCCACGAGACACAATTTGTGATCAATTAATGTTGGGCA 1261
Db 1150 AAGTAGGATCTCATGATATGCCAGTGTACAACTGTTGTTGTAATGATGGGCGATA 1209
QY 1262 CAGAGAGATCTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTATG 1321
Db 1210 CATAGGATCCAAACACGTTGGGATGATCCGATAGCTTTAAACACGAGCGTTTGAAGA 1269
QY 1322 GAAGAGAGAGAGAAAGTTGGTAGCATTTGGCATGGGAAGAGGCGTTGCCAGAGAA 1381
Db 1270 GAAGAGAGAGAGCAAAAGGCTTCTGGCGTTTGGATTAGGTAGAGAGCGTGTCTCGATCG 1329
QY 1382 CCCATGCTATGCAAGTGTGAGCTTTTCTTTGGGATTTGTTGATTCATCAATCTTTGACTGG 1441
Db 1330 GGTCTGGCCCAAGAACTGTTGGAGTAGCTCTCGGGTCAATGATACATGCTTTGATGG 1389
QY 1442 AAACGAGTAGTGAAGGAAAGCTTGTATATGACAGAGAAACAATTTGGATTCACCTTGTCAAGG 1501
Db 1390 GAGAGAGTTGGGAACTGAGAGTGTGATATGAAGAGAGGAGTTGGGAATACTGTACCCAA 1449
QY 1502 TTAATTTCAATGGAGGCCATGTGCAAGGCTCGCCCACTTGGCCACTAAATTT 1552
Db 1450 GCGATTTCTTTGAAGAGCTATTGTGCAAGGCTCGTCCATTTCTACATAAGATT 1500

US-09-938-842A-1959
 ; Sequence 1959, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRI1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1959
 ; LENGTH: 1557
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-1959

Query Match 18.9%; Score 350.6; DB 10; Length 1557;
 Best Local Similarity 56.6%; Pred. No. 1.5e-83;
 Matches 838; Conservative 0; Mismatches 589; Indels 54; Gaps 8;
 QY 95 CTGGTCTATTCTCGCGTGAAGTTGTTTTCACAGAGAGAAATGAGAAATACCA 154
 DB 31 CTGTTGTCCTATTCATCGTCAAGATGTTCTCTAGTCACTAAACCTAACAAAACTT 90
 QY 155 CCAGGTCCTCC-----TCCTCTCCATAATAGGAAACCTTAACCTCTCTCGAAGCA 208
 DB 91 CCACCGTCGCAATATATGTTTCCGATCATCGACACCTTCACCTCTCTCAAAAGCCA 150
 QY 209 ---ATCCACCGTTCTTCCAAACGATGTGAAACAGTACGGCAACGTTTCCCTCTCG 265
 DB 151 CTGCTACACGGACTCTCAGCCATCTCTCCACTCTCCGCGCCCGCTCTTTCTCTACGC 210
 QY 266 TTGGTTTCAGTCTGGCGGTGTCATCTCTCTCCACAGCATACCAAGATGCTTC--- 322
 DB 211 CTGGCTCTCGCTTCCGCTTATCATCTCATCTCCAAACGCGCAGCAAGATGCTTCTTG 270
 QY 323 ACCAAACACGACGTGCTTGGCCCAACCGGCTACCTTCTCTCTCGGAAATACATCTTC 382
 DB 271 ACCAAACACGACATGTTTACGAAATCGGCTCTGTTTATCATGCGCAAGTATGGCT 330
 QY 383 TACAACACACACCGTGGCTCTCTGCTCCACGCGAGCATGCGGCAACCTCCGCGCG 442
 DB 331 TATGACTACACCTCCATGTTACGGCTCTCTTACGGTGACCACTGGCGAAACCTCCGCT 390
 QY 443 ATCAGCGCTGGAGCTCTCCACGCGCGGTGCTCTCTCTCCGAAATCCGAGC 502
 DB 391 ATCAGCGCTTGAAGTCTCTCCAAATTCGATCTCAACGCTTCAACGCTTCAACGCTTCCGACAT 450
 QY 503 GACGAGACGAGGCTCTGATGACAGAGTTGGTGTGCGCCAAAGAACTCGAAGGAGAG 562
 DB 451 GACGAGTAAAGTCTCTCCAAAGCT-----TCACGATTATCTGTGCAACGC 501
 QY 563 TTTGCGGAGTGAAGTATGTTGATGTTTCAACGACTTAACTTACAAACATTAATGAGG 622
 DB 502 CCGCTAAAGTGGAGTTACGACAGTTACTAACGGGTCTTACCCCTAAATGTGATAATGAGA 561
 QY 623 ATGATATCGGGAGAGGTTTACGAGAGGAGTGCAGATGAAGACGTTGAGGAGCG 682
 DB 562 ATGATACCGGAAGAGATTTCTC---GAGGAAGATGATGCGGAAGGCGGGGTAAGT 618
 QY 683 AGGAGTTTCAGAGACTGTGACAGAAATGTTGGAATCATGTGGCTTGGCTTAACAGGGA 742
 DB 619 TTGAGTTCCGGAACTAGTGGCGAGATTTCTAGAGCTCTCTCGCGCGCAATCCAGCT 678

QY 743 GATCACTTGGCTTCTCCTCAGCTGGTTCGATTTTCAGAAATGTGAGAAAGCGCTTAAAGAGT 802
 DB 679 GATTTTTCGCGGCTCTACGGTGGTTTGACTACAAAGGTTTGGTCAAGAGAGCAAGAGG 738
 QY 803 ATCAGTAAGAGGTACGATTCCATCTTGAATAAGATCCCTTCATGAGAACCGTCCAGCAAT 862
 DB 739 ATTGAGAAAGATGATAGTATTATTAACAGGGTTTTTGGATGAGCATAGAGCTAACAA 798
 QY 863 GACCG-----CCAGAAATCCATGATCGATCATCTCTCAAACTCAAGAGAGCCAG 913
 DB 799 GACGATTGGAGTTTAAAGAACACCATGATTGCTCATTTTACTTGTCTCAAGAAAAAGAG 858
 QY 914 CCTCAGTACTACATGACCAAAATCATCAAGGCTTGTCTGGCCCATGCTTTTGGTGA 973
 DB 859 COTCAATACTACAGTATCAAAACCATCAAGGCTTCTTATACATGATGATGATGATGAGG 918
 QY 974 ACTGACTCATCAACTGGGACTTTAGAGTGTTCATTTATTTTAAATTTGAATCACCAGAG 1033
 DB 919 ACAGATACATCGGCTTAAACCGTGAATGGCAATGTCAAAATTTTGAATCATCCANA 978
 QY 1034 GTGTTGAAGAGGCAAGAGATGAATTGGACACTCAAGTGGGACAA-----GACCGC 1084
 DB 979 ATACTAGAGACCAAGAGCAAAACATTCATCTCAATGGAACATCAAGTAGTCTGCT 1038
 QY 1085 TTGTTAAATGATGACCTTCCAAAACCTTCCATATCTTAGGAAGATCATCTCTGAGACA 1144
 DB 1039 TTGTTGAAGAGAGAGATTTGGTAAACATGAACCTACTTTGAAATAATGTTGTGCGGAGCA 1098
 QY 1145 CTTAGTGTGTACCCCGGCCCCCAATCTTAATACCTCATGTCTTCAGAAAGATATTACA 1204
 DB 1099 CTAGACCTTTACCTGTGCGCACCTTATGTTTCTCATGTGCCATCTTCTGACTGTGTA 1158
 QY 1205 ATTGAAGGATTCATATCCACGAGACACAATTTGTGATCATTTAATGGTGGGCGATGCG 1264
 DB 1159 ATCGTGTGATTCACGTTCCAGCTGATAGATCGTGTGTTTAACTTGTGGCTATACAT 1218
 QY 1265 AGAGATCTCAGTTTGGAAATGATGCCACATGCTTTTAAACCTGAGAGGTTTGAATG- 1319
 DB 1219 CGAGATCCAAAGCGTGGGATGATCCAAACGCTATTTAAAGCCGAGAGATTTGAAGGAGT 1278
 QY 1320 -----TGGAGGAGAGAGAGAAAGTTGGTAGCATTTGGCATGGGAAGAGGCTTGC 1372
 DB 1279 GATCAGTTTGACATTAACATGTTAAATGATGCCGTTTGGTTAGGAAGAGCGGCTTGT 1338
 QY 1373 CCAGAGAACCCATGGCTATGCAAGTGTGACGTTTACTTTGGGATTTGTGATCAATGT 1432
 DB 1339 CCGGTTTGAGTTTGGCTAACCGGTTGTGGGTTTATTATTGTTTTCGATGATTCAGTGC 1398
 QY 1433 TTTGATCGAAACGAGTAAAGTGAAGAAAGCTTTGATATGACAGAGAAACAAATTGGATCACC 1492
 DB 1399 TTTGATGGGAAGCGCTCAGGAGGTGAGGTTGATATGATCTGAAGGTTCCAGGACTTAGT 1458
 QY 1493 TTGTCAGGTTAAATTCATTTGAGGCGCATGTGCAAGGCTCG 1533
 DB 1459 TTGCTTAAGCTGAACCATTTGGTTGTCAGTGCAGAACAG 1499

RESULT 8
 US-09-938-842A-1541
 ; Sequence 1541, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRI1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866

;; PRIOR FILING DATE: 2000-08-24
;; PRIOR APPLICATION NUMBER: US 60/264,647
;; PRIOR FILING DATE: 2001-01-16
;; PRIOR APPLICATION NUMBER: US 60/300,111
;; PRIOR FILING DATE: 2001-06-22
;; NUMBER OF SEQ ID NOS: 5379
;; SEQ ID NO 1541
;; LENGTH: 1548
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana

US-09-938-842A-1541

Query Match 17.8%; Score 330.4; DB 10; Length 1548;

Best Local Similarity 54.5%; Pred. No. 4.1e-78;

Matches 805; Conservative 0; Mismatches 641; Indels 30; Gaps 6;

QY	86	GTCTTTCTCTGTTCTATTCTCTCGGCGTGAAGTTGTTTTCAGAGCAAAATTTGAGA	145
DB	76	GCCTTCTCTCTTTCTATTCTCTGTTAAACATTTTGTGTAACGGCTTAGCAAGTTTAAAC	135
QY	146	AACATACCAACAGTCTCTCTCTCTCTCCATATAGGAACCTTAACCTCTCTCGAACAG	205
DB	136	CTACCAACAGTCTGCGGTAGCCCTGCTGTTGATGGACACCTTACCTCTCTGAAATC	195
QY	206	CCAATCCACCGTTTCTTCCACGCGATGCGAAACAGTACGGCAACG-----TGGTTTCC	259
DB	196	CGCTTCAACGCGCATTTCTCTCTCTCTCAATCTCTAGTGGCGCCCAATTTTCTGC	255
QY	260	CTCTGTTTGGTTCACTGCTGGCGTTGTCATCTCTCTCCACAGCATACCAAGATGC	319
DB	256	CTCGCTCTCGAAACCGTCTCACGGTTGTAGTCTCTTCTATACCTCCATCGCTGAAGAATGT	315
QY	320	TTACCAAAACACAGCGTTGCTTGGCCAAACCGGTACCTCTCTCTCGGAAATATCATC	379
DB	316	TTTACAAAACAGCATTTGTTTCGCAATAGCCAGAGTTAATCTCGGGAAACATATT	375
QY	380	TTCTACAAACACCAACCGTAGGCTCTCTGTCACACGCGAGCATCTGGCGCAACCTCCGC	439
DB	376	GAGTACAACTCCACCACTATGACTAGTGGCGCTTATGGTACCACTGGCGAAACTTGGCG	435
QY	440	CGCATACCGCGCTGGAGCTCTCTCCACGCGCGCTCCACTCTCTCGGAACTCCG	499
DB	436	CGCATCTACCTGGAGATCTTCCGATCATATAGCTCAACGGCTTCTATCTGTCGGT	495
QY	500	AGCGACGAGCAAGACGCTCTGATCAGAGGTTGCTGTCGCCAAGAACTCGAACGAGAA	559
DB	496	AAGATGAGATCCGCTCAGTTGCTAC-----TACGTTCTCCAAAACCTCGA---GACAT	546
QY	560	GAGTTTGGCGAGTGGAGATTAGTTGATGTTCAACGACTTAACCTTACAAACATATG	619
DB	547	GGGTTTGGCAAGTGGAGATGAGACATTTGCTTTTGGATTGACTATCAATAACGTTTTT	606
QY	620	AGGATGATTCGGGAGAGGTTTACGGAGGAGTGGATGAGTGAAGACGTTGAGAA	679
DB	607	AGAATGGTAGCGGAAACGATTTTATGTGAGGGCAACAGAGTGGTGGCA-665	
QY	680	CGGAGGAGTTCAGAGACTGTGCACAGAAATGTTGGAACCTCATGCGCTTGGCTTAACAAG	739
DB	666	-----GCNAGTGACATCTGATGACGAGATTGTGACCGCTGGCGCGGAATGCA	720
QY	740	GGAGATCATCTTGGCTTCTCAGTGGTTGATTTTCCAGATTTGAGAGAGCGCTTAAG	799
DB	721	GCCGACTATTTTACCAATCATCGTTGGTTTCA-----AATTTTGAAGAACGCGTCAAG	774
QY	800	AGTATCAGTAGAGTACGATTCCATCTTGAATAGATCTTCCATGAGACCGTGCCAGC	859
DB	775	AATTTAGGATTCGTATCGATAAGTTTGTGAGAGTCTTGGATGAGAAACGTCGGAT	834
QY	860	AATGACCGCCAGAAATTCATCATCGATCATCTCTCAACCTGCAAGAGACCCAGCTCAG	919
DB	835	AAAGAAAGGGTACCATCATGATAGTACCTTCTCTCTCCAGAAATCTCAGCTGAT	894
QY	920	TACTACATGACCAAAATCATCAAGGCGCTTCTCTGCGCATGCTTTTGGTGGAACTGAC	979

DB	895	TACTATACAGATGTCACTCTTAAAGGAATATATAATTTGTAATGATAATTTGCTGATCTGAA	954
QY	980	TCATCAACTGGGACTTTAGAGTGGTCAATTATCTAATTTTATTAATCACCAGAGTGTG	1039
DB	955	ACAATAGCATGGACTTTTAGAATGGCGGATGCTGAATGTGTTGAACCATCCAGAGTATTA	1014
QY	1040	AAGAAGGCAAGAGATGAATTTGACACTCAAGTGGGACAAGACCGCTTGTAAATGAGTCA	1099
DB	1015	AAGAAAGCTAGGACCGAAATCGATACGAAATCGTTTTTGACCGGTTAATGATGAAGCT	1074
QY	1100	GACCTTCCAAAACCTTCATATCTTAGGAAGATCATCTTGGAGACACTTAGGTGTACCCC	1159
DB	1075	GATACATAAAATCTGCTTATCTCCAATGGATTTGTTGGAGACCTTAGCTTTCACCCG	1134
QY	1160	CGGCGCCCAATTTCTAATACCTCATGTCTTCCAGAAATATTACAAATTTGAAGGATTCAT	1219
DB	1135	CGAGCTCCAAACAAATGTTCCATAGTAGTCAAGAGATTTGATGTTGGCAGGATATGAT	1194
QY	1220	ATCCCAACGAGACACAAATTTGTATCAATTAATGGTTGGGCGATGCAGAGAGATCCTCAGTTG	1279
DB	1195	GTTCCACGTTGGTTCGATGTTATTTGGTGAACATATGTCATGATAGACCCGCTATA	1254
QY	1280	TGGAATGATGCCACATGCTTTTAACTGAGAGTTTGAATGTTGAAGGAGGAGGAGAAAAG	1339
DB	1255	TGGGAAGACCCAGAAATGTTCAAGCCAGAGAGGTTTCAAAATTTGAAAAGCTTAAATCAAAAG	1314
QY	1340	TTGTTAGCATTTGGCATGGAAGAGGCTTCCCGAGGAGAACCCATGGCTATGCAAGT	1399
DB	1315	TTGTTGTCGTTGGGATGGACGACGAGCTGTCGCGGTATGGCTAGCTCATCGGTA	1374
QY	1400	GTGAGCTTTACTTTGGGATTTGTTGAATTTCAATGTTTGAATGAAACGAGTAAGTAGGAA	1459
DB	1375	GTGAGCTTGACTTTGGATCGATGTTTCAATGTTTGAATGCGAGAGATTTGACAAAGAA	1434
QY	1460	AGCTTG---ATATGACAGAGAACAAATTTGGATCACTTGTCAAGGTTAATTTCCATGGAG	1516
DB	1435	TATGTTGCAATAGTGAAGACAAAACTGTGGTCTTGTATGCGTCTTACGACACCGTTGCTA	1494
QY	1517	GCCATGTGCAAGCTCGGCCACTTGCCACTAAAAATTT	1552
DB	1495	GCCATGTGTAAGCTCGTCCCATTTGTCATGAGATT	1530

RESULT 9

US-09-770-149-144

; Sequence 144, Application US/09770149

; Patent No. US2002005963A1

; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kricker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; TITLE OF INVENTION: thaliana

; FILE REFERENCE: 2024 (PARA-013PRV)

; CURRENT APPLICATION NUMBER: US/09/770,149

; CURRENT FILING DATE: 2001-01-26

RESULT 11
US-10-259-165-433
; Sequence 433, Application US/10259165
; Publication No. US20030135888A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US/10/259,165
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 433:
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-259-165-433

Query Match 10.2%; Score 189; DB 12; Length 975;
Best Local Similarity 53.6%; Pred. No. 4e-40;
Matches 471; Conservative 0; Mismatches 390; Indels 18; Gaps 3;

Qy 674 GAGCAGCGAGGAGTTCAGAGACTGTGACAGAAATGTTGGAATCTATGAGTCTTGGCT 733
Db 64 GTGGAAGCCAGGAGTTTAAGCAGGTCGTGACAGAGATCATCCGACATCGCGCGGCC 123

Qy 734 AACAAAGGAGATCATCTTCTCAGGTGGTTCGATTTTCAAGATGT---GGAGAAG 790
Db 124 AACCTGTGGACTACTTTCGCGCGCTCGGTGGTTCGACGTGTTCCGCGATCATCCGACATCGCGCGGCC 183

Qy 791 CGTTAAAGATATCAGTAAGAGTACGATTCATCTTTGAATAAGATCTCTTCATGA---- 846
Db 184 ATCTCGCGCTGTAAAGCGGAGGCGGCTTCTCTCGCGCTGTATCGACGCGGCGG 243

Qy 847 -----GRACCTGCCAGCAATGACCGCCAGATTCATGATCGATCATCTCTCAAA 898
Db 244 CGGAGGTGGACGACGCGGCGAGGCGGAGAAAGAGATGATCGCGCTGCTCACT 303

Qy 899 CTGCAAGAGACCCAGCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCTGGCC 958
Db 304 CTGCAGAGACAGAGCGCGAGGTGTACACCGATTAACATGATCAGCTTAACGCGGAC 363

Qy 959 ATGCTTTTGTGGAACTGATCTATCACTGGGACTTTAGAGTGGTCAATTAATTA 1018
Db 364 TTGTTTCGAGCAGGAACAGACAACTCTCGACGATCAGAAATGGCGGATGTCGCTACTG 423

Qy 1019 TTGAATACCCAGAGGTGTTCAAGAGCGCAAGATGAATTTGGACATCAAGTGGGACAA 1078
Db 424 CTGAACACCCCGACACACTCAAGAAAGCGAAGCGGAGATCGACGATTCGTCGCGAAC 483

Qy 1079 GACCGCTTGTAAATGAGTACAGACTTCCAAAATTTCCATATCTTAGGAAGATCATCTT 1138
Db 484 TCTCGCTGATCACCAGCGGACGATGATCTCGCTCGCTACCTCCAGTGCATCGTCAGG 543

Qy 1139 GAGACACTTAGGTTGTACCCCGCGCCCAATTTCAATACCTCAATGTGTCTTCAGAGAT 1198

Db 544 GAGACGCTCCGCTGTACCCCGCGCGGATGCTCTCCCGCACGAGTCTCCCGCGAC 603
Qy 1199 ATTACAAATTGAAGGATTCAATATCCACGAGACACAAATTGTGATCATTAATGTTGGGCG 1258
Db 604 TCCAAGGTGCGCGGCTACAAATCCCGCGCGGTCGATGTTCTCATCAACGCGTACGCC 663
Qy 1259 ATGCAGAGAGATCTCTCAGTTGTGGAATGATGCCATGCTTTAAACCTGAGAGGTTGAT 1318
Db 664 ATCCACCGTGACCCGCGGTGTGGAGAGCGCGAGAGTTTCATGCCGAGAGGTTTCGAG 723
Qy 1319 GTGGAAG---GAGACGAGAAAGTTGTTAGATTTGGCATGGGAGAGGCTTGCCCA 1375
Db 724 GACGCGCGGTGCGACGCAATCTCTTGATCGCTTCGGGATGGGAGCGGAGGTTGCCCC 783
Qy 1376 GGAGAACCCATGGCTATGCAAAAGTGTCTGCTTTTCTTTGATTTGATTTCAATGTTT 1435
Db 784 GCGAGACGCTGGCGCTGCGACAGTGGGTTGGTGGGACGCTGATCCAGTGTTC 843
Qy 1436 GACTGGAAACGAGTAAGTGAGGAAAGCTTGTATATGACAGAGACAAATTTGATTCACCTTG 1495
Db 844 GACTGGAGAGGTCGACGCGCTGAGGTTCGACATGACTGAAGGTGGCGGCTCACCATC 903
Qy 1496 TCAAGGTTAAATTCATTTGGNGCCCATGTCRAGGCTCGC 1534
Db 904 CCCAAGGTGCTGCCCTTGGAGGCCATGTGACGCGCGC 942

RESULT 12
US-10-259-165-101
; Sequence 101, Application US/10259165
; Publication No. US20030135888A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 101:
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-259-165-101

Query Match 10.2%; Score 189; DB 12; Length 978;
Best Local Similarity 53.6%; Pred. No. 4e-40;
Matches 471; Conservative 0; Mismatches 390; Indels 18; Gaps 3;

Qy 674 GAGGAAGCGAGGAGTTCAGAGACTGTGACAGAAATGTTGGAATCTATGAGTCTTGGCT 733
Db 64 GTGGAAGCCAGGAGTTTAAGCAGGTCGTGACAGAGATCATCCGACATCGCGCGGCC 123

Qy 734 AACAAAGGAGATCATCTTCTCAGGTGGTTCGATTTTCAAGATGT---GGAGAAG 790
Db 124 AACCTGTGGACTACTTTCGCGCGCTCGGTGGTTCGACGTGTTCCGCGATCATCCGACATCGCGCGGCC 183

Qy 791 CGTTAAAGATATCAGTAAGAGTACGATTCATCTTTGAATAAGATCTCTTCATGA---- 846
Db 184 ATCTCGCGCTGTAAAGCGGAGGCGGCTTCTCTCGCGCTGTATCGACGCGGCGG 243

Qy 847 -----GRACCTGCCAGCAATGACCGCCAGATTCATGATCGATCATCTCTCAAA 898
Db 244 CGGAGGTGGACGACGCGGCGAGGCGGAGAAAGAGATGATCGCGCTGCTCACT 303

Qy 899 CTGCAAGAGACCCAGCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCTGGCC 958
Db 304 CTGCAGAGACAGAGCGCGAGGTGTACACCGATTAACATGATCAGCTTAACGCGGAC 363

Qy 959 ATGCTTTTGTGGAACTGATCTATCACTGGGACTTTAGAGTGGTCAATTAATTA 1018
Db 364 TTGTTTCGAGCAGGAACAGACAACTCTCGACGATCAGAAATGGCGGATGTCGCTACTG 423

Qy 1019 TTGAATACCCAGAGGTGTTCAAGAGCGCAAGATGAATTTGGACATCAAGTGGGACAA 1078
Db 424 CTGAACACCCCGACACACTCAAGAAAGCGAAGCGGAGATCGACGATTCGTCGCGAAC 483

Qy 1079 GACCGCTTGTAAATGAGTACAGACTTCCAAAATTTCCATATCTTAGGAAGATCATCTT 1138
Db 484 TCTCGCTGATCACCAGCGGACGATGATCTCGCTCGCTACCTCCAGTGCATCGTCAGG 543

Qy 1139 GAGACACTTAGGTTGTACCCCGCGCCCAATTTCAATACCTCAATGTGTCTTCAGAGAT 1198

```

Db 124 AACCTGGGACTTTCGCCGCGCTCCGGTGGTTCGACGTGTTCGGCGTCAGGAGAG 183
Qy 791 CGTTAAAGAGTATCAGTAAGAGGTACGATTCCATCTTGAATAAGATTCCTTCATGA ---- 846
Db 184 ATCTCGCCGCTGTAAGACCGGAGGACGCGTTCCTTCGCCGCTGATCGACGCGGAGCG 243
Qy 847 -----GAACCGTGCCAGCAATGCCGCCAGATTCGATGATCGATCATCTCTCAAA 898
Db 244 CGAGGCTGACGACGCGGACGAGGCGGAGAGAGCATGATCGCGGTGCTGCTCAT 303
Qy 899 CTGCAAGAGACCCAGCGCTCAGTACTACACTGACCAAAATCATCAAGGCGCTTGTCTGGGC 958
Db 304 CTGCAAGAGACAGCGCGGAGGTACACCGATAACATGATCACAGCTCTAAGCGGAGAC 363
Qy 959 ATGCTTTTGGTGGAACTGACTCATCAATCGGAGCTTTAGAGTGGTCAATTAATTA 1018
Db 364 TTGTTGGAGAGGAACAGAGACAACCTCGACGACATCAGAAATGGGCGATGTCGCTACTG 423
Qy 1019 TTGAATCACCAGAGGTGTTGAAGAGGCAAGAGATGAATGGACACTCAAGTGGGACAA 1078
Db 424 CTGAACCCACCCGACACACTCAAGAAAGCCGACGAGATCGAGCATCCGTCGGCAAC 483
Qy 1079 GACCGGTTGTTAAATGAGTCAGACCTTCCAAAACCTTCCATATCTTAGGAAGATCATCCTT 1138
Db 484 TCTCGCTGATACCCGCGACGAGTGACTCGCTCGGTACCTCCAGTGCACTGTCAGG 543
Qy 1139 GAGACACTTAGTTGATCCCGCGGCGCCAAATCTAATACCTCATGTGCTTCAGAAAGAT 1198
Db 544 GAGACGCTCGCGCTGTACCCGCGCGCGCGATGCTCTCCCGCACGAGTCTCCGCGGAC 603
Qy 1199 ATTCAATTCAAGATTCATATCCACGAGACACAAATGTCATATTAATGTTGGGCG 1258
Db 604 TGAAGGTGCGCGGCTACAAATCCCGCGCGGTCGATGTTGCTCATCAACGCGTACGCC 663
Qy 1259 ATCAGAGAGATCTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGAT 1318
Db 664 ATCCACGTGACCGCGGCTGTGGAGGACCGGAGAGTTTCATGCGGAGAGGTTGCGAG 723
Qy 1319 GTGGAAG--GAGAGGAGAAAAAGTTGGTAGCAATTTGGCATGGGAAGAGGCTTGCCCA 1375
Db 724 GACGGGGGTGCGACGCGCAATCTCTTGATGCCGTTGGGATGGGAGCGGAGGTGCCCG 783
Qy 1376 GAGAACCCATGCTATGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1435
Db 784 GCGGAGACGCTGGCGTGCACAGTGGGTTGGTGGGCGACGCTGATCCAGTGTCTTC 843
Qy 1436 GACTGGAACAGTAGTGAAGGAAAGCTTGATATGACAGAGAAACAAATTCGATCACTTG 1495
Db 844 GACTGGAGAGGTGCGACGCGGTGGAGTGCATGATGAGGTGGCGGCTCACCATC 903
Qy 1496 TCAAGGTAAATTCATGGAGGCGCATGTGCAAGGCTGCG 1534
Db 904 CCCAAGTGTGCGGTTGGAGGCGCATGTGCGAGCGCGCG 942

```

RESULT 13

```

US-10-174-693-221
; Sequence 221, Application US/10174693
; Publication No. US20030131373A1
; GENERAL INFORMATION:
; APPLICANT: Blosberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c5
; CURRENT APPLICATION NUMBER: US/10/174,693
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09

```

```

; PRIOR APPLICATION NUMBER: US 09/615,192
; PRIOR FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 407
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 221
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-174-693-221

```

```

Query Match 9.8%; Score 182.2; DB 15; Length 447;
Best Local Similarity 64.9%; Pred. No. 1.6e-38;
Matches 290; Conservative 0; Mismatches 148; Indels 9; Gaps 1;

Qy 1038 TGAAGAGCGCAAGATGAATGGACACTCAAGTGGGCAAGACCGCTTGTAAATGACT 1097
Db 1 TAAATGAAGCGCCCAAGATGAGATTCATCTATGATTTGGCCATGATAGTTTGTAGAAGAT 60
Qy 1098 CAGACCTTCCAAACCTTCATATCTTAGAAGATCATCTTTGAGACACTTTAGTGTGTACC 1157
Db 61 CGGATGTTTCAAAACTACCTTACCTTCAGTGCATTAATCTTGAGACCTTTCGACTAAACA 120
Qy 1158 CCCCAGCGCCCAATCTAATACCTCATGTGTCTTCCAGAAAGATTAACAATGAAGATTCA 1217
Db 121 CGACGCGCACCTTCTCTCCACACGCGTCATCGCTGATTCGACTATAGGAGGATCT 180
Qy 1218 ATATCCCAAGACACAAATGTGATCATTAATGTTGGGGCATGAGAGAGATTCCTCAGT 1277
Db 181 TCGTCCCAAGCGACACTATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 240
Qy 1278 TGTGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTATGTAAGAGAGGAGAGAAA 1337
Db 241 TGTGGAGGATGTCATTTGAGCTTCAAGCCTGAAAGGTTTCAGGCGCAATGCGACGAAAGC 300
Qy 1338 A-----GTGTGTAGCATTTGGCATGGGAAGAGGCTTGGCCAGGAGAACCCATCG 1388
Db 301 AACAAAGCTACTATTGCTTTTGGACGAGGAGGAGGATGCGGCGGAGGATGCG 360
Qy 1389 CTATGCAAGTGTGAGCTTTTACTTTGGGATGTTGATTCATTAATGATTTTGTGCTGAAACGAG 1448
Db 361 CTCATCGGTGTCATGGGGTGGACGTTGGGCTTGTGATTCAGTGTGTTGATTCGAAAGAG 420
Qy 1449 TAAGTGAGGAAAGCTTGCATATGACAG 1475
Db 421 TAAAGCAAGAGAGATTGACATGACGG 447

```

RESULT 14

```

US-10-356-153-55
; Sequence 55, Application US/10356153
; Publication No. US20030166176A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/356,153
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US/10/142,231
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Taxus cuspidata
US-10-356-153-55

```

```

Query Match 9.8%; Score 181.6; DB 12; Length 1545;
Best Local Similarity 48.5%; Pred. No. 5.4e-38;
Matches 727; Conservative 0; Mismatches 714; Indels 57; Gaps 6;

```



```

Db      | 340 GTAGGATGATGACATCAAAATCTTCAGTATGGCTCTTACGGTTCCTACTGGCGAAACCTT 399
Qy      | 437 CGCGCGATCACCGCCCTGGACGTCCTCTCCAGCGCGGTCCACTCTCTCTCCGGAATC 496
Db      | 400 CGGAAATGTGCAGCATCCAGATCTCTCTGCAACAGAAATGGACTCTCTTCAGACATC 459
Qy      | 497 CGGAGCGAGAGACGAGGCTGTGATGAGAGGTTGGTCTGGCCAGAACTCGAACGAG 556
Db      | 460 CGCGTAGAGGAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 510
Qy      | 557 GAAGAGTTTGGCGGAGTGGAGATTAGTTCGATCTCAACGACTTAACTTAACTTAACTTAA 616
Db      | 511 CGAGAGGACACTCCAGTCAACATGAAGCGAGGCTCTCTGATCTCAGCTTGTAGTATCATC 570
Qy      | 617 ATGAGGATGATATCGGGGAGAGGTTTACGGAGGAGAGAGTGAATGAAGAACGTTGAG 676
Db      | 571 CTCGGTATGGTTCGCAACAGAAATTAACAGACCTG---TTTATTCGGAGGAATACGAA 627
Qy      | 677 GAAGCGAGGAGTTACAGAGACTGTGACAGAAATGTTGGAATCATGGGCTTGGCTAAC 736
Db      | 628 GAAGCGGATCAATTTAAACAGATGATAAAACAGTCTGTGTTCTTACTTGGAGCAITTTGAG 687
Qy      | 737 AAGGAGATCACTTGGCTTTCTCTCAGGTGTTTCGATTTTCAGAACTGTGAGAGGCGCTTA 796
Db      | 688 GTTGGAGATTTCTCGCGTTTCTCAAGTGGCTTGAATCTCAGGGTTTCTAGCTGCTATG 747
Qy      | 797 AAGAGTATCAGTAAGAGGTACGATTCATCTTGAATAAGATCCTTCATGAGAACCGTGCC 856
Db      | 748 AAAAACTGCAGCAGAAAGAGATGCTTTATGACAGAAATGGTGATTGATCACCGTGAG 807
Qy      | 857 AGCAATGACCGCCAGAAATTC-----CATGATGATCATCTCTCAAACTGCAA 904
Db      | 808 AAGAGAGGAGAGTTCGATGCAAAATGCAAGACTTAAATGATGTTCTCTCTGCAACA 867
Qy      | 905 GAGACCCAGCTCAGTACTAC-----ACTGACCAATCATCAAGGCTTGTCTCTGGCC 958
Db      | 868 GACAACCATGAATTCAGTCCGATAGTAACGAGATGTTGTGAAGCCACCGCCCTTACA 927
Qy      | 959 ATGCTTTTGGTGGAACTGACTCATCAACTGGGACTTTAGAGTGGTCAATTCATTAATTTA 1018
Db      | 928 ATGCTGAAACGAGGTACAGATACATCTCGGTGACCATCGAATGGGCAATGGCGGCTCTG 987
Qy      | 1019 TTCAATCACCCAGAGGTTTGAAGAGAGGCAAGAGATGAATTTGACACTCAAGTGGACAA 1078
Db      | 988 ATGACGACCCCTCATATTTTGAAGAAAGCCAGCAGGAGCTCGACACGATATCGGACGC 1047
Qy      | 1079 GACCGCTTGTAAATGATGATGAGCTTCCAAACTTCCATATCTTAGGAGATCATCCTT 1138
Db      | 1048 AGCCGATTTAGAGGAGCAGATCTGACGAGCTGAAATATTTGAGGCAATTTGTGAA 1107
Qy      | 1139 GAGACACTTATGTTGTACCCCGGCGCCCAATTTCTAATACCTCATGTGTCTCAGAAAT 1198
Db      | 1108 GAAACGTTAGGCTATATCCAGCGGACCTCTCTTAGTTCCTCAGAACCATTTAGGAT 1167
Qy      | 1199 ATTACAAATTCAGAAATTCATATCCACAGAGACACAAATGTGATCATTAATGTTGGGC 1258
Db      | 1168 TGCACTGTTGGAGGTTACCATGTCTCGCAGGAAACGCACTGATTTGTAATGCTTGGGCA 1227
Qy      | 1259 ATCAGAGAGATCTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGT----- 1313
Db      | 1228 ATTCACAGACCCCGGAGTGTGGGAACCGGCCACCGTGTTCGATCCTGAAACGGTTTGTG 1287
Qy      | 1314 -----TTGATGTGAAGGAGAGAGAGAAAGTTGGTAGCATTTGGCATG 1357
Db      | 1288 AAGAGCGGAAAGAGGTTGACGTAAAGAGGGGGAGTTTGAATTTGATTCCGTTGGTTCA 1347
Qy      | 1358 GGAAGAGGCTTGCAGAGAGAACCCATGGCTATGCAAGTGTGACGTTTACTTTGGGA 1417
Db      | 1348 GGGAGAAGAAATGTGTCCGGGCATGAGTCTGGCATTTGAGTGTGTTACGTATACGCTGGGG 1407
Qy      | 1418 TTGTTGATTCATGTTTGTGACTGGAAACGAGTAAGTGAGGAAAGCTTGTATATCACAGAG 1477

```

```

Db      | 1408 AGGCTGCTGCAGAGCTTCGAGTGTCTCTTCCAGNAGGTATGATAATTTGACATGACGAA 1467
Qy      | 1478 AACAAATTGGATCACCCTTGTCAAGGTTAATTCATTGGAGGCCATGTGCAAGGCTCGCC 1535
Db      | 1468 GGTTTGGGACTTCACAATGCCCAAGCAGTTCCTGTTGGAGACCATTTATCAAAACCTCGCC 1525

```

Search completed: October 21, 2003, 08:02:41
 Job time : 1663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 02:43:04 ; Search time 133 Seconds
(without alignments)

6169.408 Million cell updates/sec

Title: US-09-868-546A-1

Perfect score: 1859

Sequence: 1 gaaacacgacgacagacga.....aaataattcttactttc 1859

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	328.4	17.7	1698	4	US-09-059-769-20
2	182.2	9.8	447	4	US-09-615-192A-221
3	143.4	7.7	494	4	US-09-615-192A-222
4	131.8	7.1	1704	3	US-08-948-564-17
5	128.6	6.9	1821	1	US-08-313-075A-37
6	128.6	6.9	1824	3	US-08-606-505B-1
7	128.6	6.9	1824	3	US-09-616-990-1
8	127.2	6.8	622	4	US-08-975-316-17
9	127.2	6.8	622	4	US-09-615-192A-17
10	120.8	6.5	1812	1	US-07-912-900-28
11	120.8	6.5	1812	1	US-08-285-309-28
12	120.8	6.5	1812	2	US-08-502-046-28
13	118.2	6.4	1755	1	US-07-912-900-29
14	118.2	6.4	1755	1	US-08-285-309-29
15	118.2	6.4	1755	2	US-08-502-046-29
16	117	6.3	2174	3	US-08-606-505B-63
17	117	6.3	2174	3	US-09-616-990-63
18	116.6	6.3	1691	3	US-08-948-564-3
19	114.6	6.2	1929	4	US-09-380-420C-1
20	110.4	5.9	1806	4	US-09-351-229-3
21	109.4	5.9	1893	1	US-08-532-065B-1
22	108.8	5.9	1708	3	US-08-391-677-1
23	108.8	5.9	1761	3	US-09-033-055A-3
24	104	5.6	1824	3	US-08-948-564-13
25	103.2	5.6	1788	3	US-08-948-564-9
26	102.8	5.5	1657	3	US-08-948-564-11
27	101	5.4	1781	4	US-09-499-302A-1

Sequence 2, Appli
Sequence 1, Appli
Sequence 29, Appli
Sequence 24, Appli
Sequence 24, Appli
Sequence 24, Appli
Sequence 201, App
Sequence 67, Appli
Sequence 8, Appli
Sequence 3, Appli
Sequence 69, Appli
Sequence 231, App
Sequence 1, Appli
Sequence 64, Appli
Sequence 64, Appli
Sequence 203, App
Sequence 5, Appli
Sequence 7, Appli

28 99.6 5.4 1634 4 US-09-126-420A-2
29 98.8 5.3 1737 4 US-09-126-420A-1
30 96.8 5.2 1665 1 US-08-313-075A-29
31 93.6 5.0 1666 1 US-07-912-900-24
32 93.6 5.0 1666 1 US-08-285-309-24
33 93.6 5.0 1666 2 US-08-502-046-24
34 91 4.9 993 4 US-09-615-192A-201
35 91 4.9 1665 3 US-09-292-768-67
36 89.4 4.8 1665 3 US-08-881-784-8
37 89.4 4.8 1665 3 US-09-292-768-3
38 89.4 4.8 1665 3 US-09-292-768-69
39 89 4.8 381 4 US-09-615-192A-231
40 88.6 4.8 1838 3 US-08-948-564-1
41 88.2 4.7 1927 3 US-08-606-505B-64
42 88.2 4.7 1927 3 US-09-616-990-64
43 87.4 4.7 432 4 US-09-615-192A-203
44 84.2 4.5 1515 3 US-09-292-768-5
45 83 4.5 1611 3 US-08-948-564-7

ALIGNMENTS

RESULT 1

US-09-059-769-20
; Sequence 20, Application US/09059769
; Patent No. 6329518
; GENERAL INFORMATION:
; APPLICANT: Green, Allan
; APPLICANT: Singh, Surinder
; APPLICANT: Lenman, Marit
; APPLICANT: Stymne, Sten
; TITLE OF INVENTION: Plant Fatty Acid Epoxigenase Genes and Uses
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,769
; FILING DATE: April 14, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P06223
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P06226
; FILING DATE: 15-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/043706
; FILING DATE: 16-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/050403
; FILING DATE: 20-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1698 base pairs
; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8..1504
US-09-059-769-20

```

Query Match 17.7%; Score 328.4; DB 4; Length 1698;
Best Local Similarity 54.5%; Pred. No. 5.2e-76;
Matches 834; Conservative 0; Mismatches 626; Indels 70; Gaps 6;

Qy	89	CTTTTCTGGTTC	TAATTCCTCGGCGTGAAGTTGTTTTC	AAACGAGAAATTCGAGAAC	148
Db	44	CTCTTCTCTCT	TTAAATTAAGTTTCAATTTTTTAAAC	TCAAAACCCCCAAAAC	103
Qy	149	ATACCACGAGT	CTCTCTCTTCCCATTAATAGAAACCTT	TAACCTCTCTCGAACGCCA	208
Db	104	CTCCCCCATC	ACACCATCTTTTCCAAATCACCGGCCAT	CTCCATCTCTCTTAACAACCA	163
Qy	209	ATCCACGGTTT	TTCTTCCAACGATGTGAAACAGTACGG	CAACGTGGTGTTCCTCTGGTTC	268
Db	164	ATCCACAGAA	CTCTCCACCAATCGCCACCAAGTACGG	GGACATCTTATTCTCTCCGATTC	223
Qy	269	GGTTACAGT	CTGGCGTTGTCACTCTCTCCAAACAGAT	ACCAAGAAATGCTTCACCAAA	328
Db	224	GGAAACAAA	AGTCTTAGTCACTCTCTCTCCCGCGT	ACAAAGAAATGTTTCACTATA	283
Qy	329	CACGAGCTG	CTTGGCCAAACGGCTACTTCTCTCGG	AAATACATCTTCTACAAC	388
Db	284	AACGACATCA	TTTTCGCTAACCGCCCAACAAATCT	CGCGGGAAGCACTCAATTACAT	343
Qy	389	AACACCA	CCGTAGTCTCTGCTCCACGGGAGCACT	GGGCAACCTCCGCGCATCACC	448
Db	344	TCCACCA	CCATGGATTCGCTCTATGGCGATCA	CTGGCGTCACTCCGACGACTCACA	403
Qy	449	GCCCTGAG	CTCTCTCAACGACGGGTTCACCTCTT	CTCCGGAATCCGGAGCGACGAG	508
Db	404	ACAATTAG	CTCTCTCTGCAAAATCGTGTGGCA	TGTTTTCCGGGTTCGGGCGCGATGA	463
Qy	509	ACGAAGCGT	CTGATGCAGAGTTCGTGTCGCCAAGA	ACTCGAACGAGGAAGTGTTCG	568
Db	464	AGTACACT	TTTTATCAACAGTGTGTTCCAG-----	GAATTCGGATTCGGGAAG	514
Qy	569	CGAGTGAG	ATTAAGTTCGATGTTTCAACGACTTAA	CTTACAAACAATAATCAGGATGATA	628
Db	515	ATAGTA	ACTTTGACATCGAAACTGATGGAGCT	TACACTGATATACTAATGAGAAATGCT	574
Qy	629	TCGGGGA	AGAGTTTATCGGAGAGAGATGAGAA	ACGTTGAGGAAGCGAGGAG	688
Db	575	GCCGAAA	ACGGTTTATCGGGAAGA-----	AGTGAAGGATGAAGAAGTGCAGTTG	625
Qy	689	TTCAGAGAG	ACTGTGCAGAAATCTTGGAACTCATGG	-----CTTGGCTTAACAGGGAGAT	745
Db	626	TTGAGGAT	CTTATGAAGAAATGGAGCGCTCCG	GGGGAATTAACCGTGAACCGAGAT	685
Qy	746	CACTTGCT	TTCTCCAGTGGTTCGATTTTCAGAA	ATGTGGAGAGCGCTTAAAGAGTATC	805
Db	686	TAITTTCC	AGTATTTCAGTGGATTGATACAGG	AGTAAAGAGAGATGAGNAACCTG	745
Qy	806	AGTAAGAG	GTACGATTCATCTTGAATAAGATCT	TCAATGAGAACCGTGCAGCAATGAC	865
Db	746	ATGAAGAA	AAATGGACGGGTCTTTCGAAAAT	CTCATTTGATGAACACCGAAACACG	805
Qy	866	CGCAGAA	T-----	-----TCCATGATCGATCAT	889
Db	806	TGGATCA	ATCAAGTTCGAGCAACTCGGAC	AAAAAGAGAACTTGGACATCGGTAG	865
Qy	890	CTCTCA	AACTGCAAGAGACCCGCTCAGTACT	ACACTGACCAAAATCATCAAGGCGTT	949
Db	866	ATGTTG	AAATCTTAAAAAGACAAACCTG	ACTTCTTACACTGATCTAATCTATCA	925
Qy	950	GCTCTG	CCCATGCTTTTTTGGTGGAACTG	ACTCATCAACTGGCACTTTAGAGT	1009

Query Match	7.1%	Score 131.8	DB 3	Length 1704	
Best Local Similarity	45.8%	Pred. No. 1.1e-24			
Matches 546	Conservative	0	Mismatches 642	Indels 9	Gaps 2
QY	146	AACATACCACGAGTCTCTCTCTCTCCCATATATAGGAAACCTTAACTCTCTCGAACAG	205		
DB	116	AAGTCCCTCGGGTCCA CGGCCCTGGCCGGTAGTCGGTAACCTCTACGACATAAAACCC	175		
QY	206	CCAATCCACGGTTTCTTCCAACGCATGTGGAACAGTACGGCAACGTGGTTTCCCTCTCG	265		
DB	176	GTCCGCTTCGGTGTCTTCGGCGAGTGGCGCAGTCTTACGGCCCCATAATATCGTGTGG	235		
QY	266	TTCCGTTACAGTCTGGCCGTTGTCATCTCTCTCCACAGCATACCAAGAATCTTCAACC	325		
DB	236	TTCCGTTTCGACCCTAAACGTCATCGTTTCGAACTCCGAGCTGGCGAAGGAGTGTGAAG	295		
QY	326	AAACACGACGTTTCCCTTTGGCCAAACCGCTACCTTCTCTCTCGGAAATAATCATCTTCTAC	385		
DB	296	GAGCACGATCAGCTGTGGCGGACCGCCACCGGAGCGGTTCGGCGGAAGTTTCAGCCGC	355		
QY	386	AACAAACACACGTAGGCTCTGCTCCACGGCGAGCACTGGCGCAACCTCCCGCGCATC	445		
DB	356	GACGGGAAGGATCTAAATTTGGGCGCATTAATGGGCGCAGTACGTGAAGTGTGAGGAAGTT	415		
QY	446	ACGCCCTGACGTCCTCTCACGACGCGGTCCACTCTCTCTCCGAAATCCCGAGCGAC	505		
DB	416	TGCACGCTCGAGCTTTCTCGCGAAGGCGCTCGAGCCCTGAGGCCCATTAGGAGGAC	475		
QY	506	GAGACGAAGCGTCTGATGACAGAGTTGGTCTCGGCCAAGAACTCGAACGAGGAAGTTT	565		
DB	476	GAGTCACTCCATCTGTTGATCTCGTTTACAATCACTGCACACGACCTGAAAATTTGGGG	535		
QY	566	GC CGGAGTGAGATTAAGTTCGATGTTCAACGACTTAACCTTACAACAACTAATGAGGATG	625		
DB	536	AAAGGAATATTGTTGAGGAAGCATCTGGGGTTGTGGCAITCAACAACATAACGAGTTG	595		
QY	626	ATATCGGGGAAGAGTTTACGGAGAGGAGAGTGAGATGAAGAACGTTGAGGAAGCGAGG	685		
DB	596	GCATTTGGGAAAAGATTGTGAACTCAGAAGGTGTGATG-----GATGACGAAGGAGTA	649		
QY	686	GAGTTCAGAGAGACTGTGACAGAAATGTTGGAACTCATGGCTTGGCTAACAGGGAGAT	745		
DB	650	GAATTTCAAGGCCAATTTGSGAAATGGGTTAAAGCTAGGAGCATCTTAGCCATGGCGAAA	709		
QY	746	CAC TTGCTTTCTCTCAGTGGTTCGATTTTCAGAAATGTGGAAGAGCGCTTAAAGAGTATC	805		
DB	710	CACATCCCTTGGCTTCGCTGGATGTTCCCACTGGAAGAAGAGCTTTTGCCAAGCATGGA	769		
QY	806	AGTAAGAGGTACGNATTCATCTTGAATAAGATCCTTCATGAGAACCGTCCAGCAATGAC	865		
DB	770	GC CGCGCGCAGCCGCACTCAACAGAGCCATCATGCCAGAGCACACTGAAGCACCGCAGAAA	829		
QY	866	CGCCAGAAATTCATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCCCTCAGTACTAC	925		
DB	830	TCTGGTGGTGCCAGCAACATTTTGTGATGCCCTCTCTCATTTGCAAGACAATATGAC	889		
QY	926	---ACTGACCAAATCATCAAAGGCCCTTCTCTGCGCCATGCTTTTGTGTGGAAC TGACTCA	982		
DB	890	CTTAGTGAAGACACCATCATTTGTTCTCTTTGGGATATGATCAAGCAGAGGATGGAACA	949		
QY	983	TCAACTGGGACTTTAGAGTGGTCATTATCTAATTTATTTGATCACCACAGGTTGTTGAAG	1042		
DB	950	ACTGCAATTTTCAGTTGAGTGGGCCATGGCTGAGTTGATGAAGAAACCAAGGGTGCAACAA	1009		
QY	1043	AAGGACAGAGATGAATTTGGACACTCAAGTGGGAACAAGCCGCTTGTTTAAATGAGTCAAG	1102		
DB	1010	AAGGTCGAAGGAGCTAGACAGGGTAATTTGGGCTTGAAGGGTGTATCATCTGAAGCAGAC	1069		
QY	1103	CTTCCAAAACTTCCATATCTTAGGAAGATCATCTTGAGACACTTATGAGTTGTACCCCGG	1162		
DB	1070	TTCTCAATCTCCCTTACCTCAATGTGTGACCAAGAGCAATGAGGCTTCAACCCACCA	1129		

Qy	1163	GCCCGCAATCTAATAC	TGCTGCTT	CAGAGATATTAC	ATTCAATTC	GAAGGATTC	CAATATC	1222
Db	1130	ACCCCACTAATGCTCC	CAACCGTCC	AATGCCAATG	CCCAATGT	CAAGTTG	GGGCTATG	1189
Qy	1223	CCAGAGACACAAT	TGTGATCA	TAAATG	TGGGGCATG	CAGAGAGAT	CTCTCAGTTGTGG	1282
Db	1190	CCCAAAAGGTTCCAA	TGTGATGTGA	ATGTGTGG	GGCGGTGG	CCGCCAC	CGGCCGTGTGG	1249
Qy	1283	AATGATGCCACATG	CTTTAAAC	CTGAGAGG	TTTGATGTG	GAAGAGG	AGAGAGAAAAAG	1339
Db	1250	AAGGATCCATTTGG	AGTTTCGAC	CCGAAAG	GGTTCCTTTG	GAGGAGGAT	GTAGACATGAAG	1306

RESULT 5
US-08-313-075A-37
; Sequence 37, Application US/08313075A
; Patent No. 5639870
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwina C.
; APPLICANT: Tanaka, Yoshikazu
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Preeser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,075A
; FILING DATE: 30-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 1538/92
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 6698/93
; FILING DATE: 07-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU93/00127
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9433
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1812 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 126..1643
US-08-313-075A-37

623 ATGATATCGGGAAGAGGTTTACGGAGAGAGAGTGAAGAAAGTTGAGAGCG 682
670 ATATATCGGACAGATGATGCTAAACAAAGAGTATTTG-TAGATAAAGGTTGTTGAGTA 728
683 AGGAGATTACAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGGCTAAACAGGA 742
729 AATGAATTTAAGACATGTTGTTAGAGTTAATGACATAGCAGGATTTTCAACATTTGTT 788
743 GATCACTTGCCTTCTCAGTGGTTCGATTTTTCAGAAATGTTGGAAGCGCTTTAAAGAGT 802
789 GATTTATTTCTTGTAGTTGGATGATTTTACAAAGGATAGAAAACGAATGAACGT 848
803 ATCAGTAAGAGTACGATTCATCTTGAATAGATCTTTCATGAGAACCGTGGCA----- 857
849 TTACATAAGAAGTTTGTATGCTTTATGACAAAAGATGTTTGTATGAACACAAAGCAACTACC 908
858 -----GCAATGACCGCAGAAATCCATGATCGATCTCTCCTCAAACTGCAACAGAGACC 910
909 TATGAACGTAAAGGGAACAGATTTCTTGATGTTGTTATGGAATAGGGGACAAATCTT 968
911 CAGCCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCTGGCCATGCTTTTGGT 970
969 GAAGGAGAAAGACTCAGTACAAACCAACATCAAGACACTTTTCTGTAATTTGTTTACAGCT 1028
971 GGAAGTACTATCACTGGGACTTTAGAGTGGTCAATCTAATTTATTTATGATCACCACCA 1030
1029 GGTACGGACACTTCTTCTAGTGAATAGAAATGGGACCTTGCAGAAATGATGAAGAACCT 1088
1031 GAGGTGTTCAAGAAGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTTGTA 1090
1089 GCCATTTTGAAGAAAGCAAGCAGAAATGGATCAAGTCAATTTGGAAAGAAATAGGGCTTTA 1148
1091 AATGAGTACAGCTTCCAAAACCTTCATATCTTAGGAAGATCATCTTTGAGACACTTAGG 1150
1149 CTGCAATCGATATCCAAATCTCCCTTACCTCCGAGCAATTTGCAAGAAACATTTTCA 1208
1151 TTGTACCCCGCCGCAATCTAATACCTCATGTCTTTCAGAGATATTAATTAAGAA 1210
1209 AAACACCTTCTACACCAATTAATCTTCTAGGATCTCGAACCAACCAATGATGCTGAT 1268
1211 GGATTCATATCCACGAGACACAATTTGATCATTAATGTTGGGCAATGAGAGAT 1270
1269 GGTATTATACCAAAACACTAGGCTTAGTTTAAATATGGAATAGGCAATGGAAGAT 1328
1271 CCTCAGTTTGGAAATGATCCCATGCTTTTAAACCTGAGAGT----- 1313
1329 CCCAAGTTTGGAAATCCACTAGATTTTAAATCCGAAAGATTTCTTGAGTGAAGAAAC 1388
1314 -----TTGATGTGAAGAGAGGAGAAAGTTGGTAGCATTTGGCAATGGAAGAGG 1366
1389 TCCAGATTTGATCCTCGAGGGAACGATTTTGAATTTGATACCAATTTGGTCTGGAAGAGA 1448
1367 GCTTGCCAGGAGAACCCATGGCTATGCAAGTGTGAGCTTTACTTTGGGATTTGAT 1426
1449 ATTGTGAGGACACAGATGGAAATTTGAATGTTGGAATATATATTAGGAATTTGGTT 1508
1427 CAATGTTTGTACTGGAACAGAGTAAGTGAAG 1457
1509 CATTCTTTGATTGGAAATTAACCAAGTGAAG 1539

RESULT 6

US-08-606-505B-1
; Sequence 1, Application US/08606505B
; Patent No. 6114601
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: KIKUCHI, Shigeto
; APPLICANT: SHIMADA, Yukihisa
; APPLICANT: OHBAYASHI, Masaya
; APPLICANT: SHIMADA, Ritsuko
; APPLICANT: OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES

NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1824 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Petunia hybrida
STRAIN: Falcon Blue
FEATURE:
NAME/KEY: CDS
LOCATION: 116 to 1633
IDENTIFICATION METHOD: by experiment
US-08-606-505B-1

Query Match 6.9%; Score 128.6; DB 3; Length 1824;

Best Local Similarity 49.9%; Pred. No. 8e-24;
Matches 435; Conservative 0; Mismatches 399; Indels 37; Gaps 3;

QY 623 ATGATATCGGGAAGAGGTTTACGGAGAGAGTGAAGAAAGTTGAGAGCG 682
DB 660 ATATGATCGGACAGTGTGCTAAAGCAAAAGAGTATTTG-TAGATAAAGGTTGTTGAGTA 718
QY 683 AGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGGCTTAACAGGA 742
DB 719 AATGAATTTAAGGACATGTTGTTAGAGTTAATGACAATAGCAGGGGATTTTCAACATTTGT 778
QY 743 GATCACTTGCCTTCTCAGGTTGGTTCGATTTTCAGAAATGTTGGAAGCGCTTTAAAGAGT 802
DB 779 GATTTATCTCTGTTTGTAGCTTGGATGATTTTCAAGGATAGAAAACGAATGAACGT 838
QY 803 ATCAGTAAGAGTACGATTTCCATCTTGAATAGATCTTTCATGAGAACCGTGGCA----- 857
DB 839 TTACATAAGAAGTTTGTATGCTTTTATGACAAAAGATGTTTGTATGAACACAAAGCAACTACC 898
QY 858 -----GCAATGACCGCAGAAATCCATGATCGATCTCTCCTCAAACTGCAAGAGACC 910
DB 899 TATGAACGTAAAGGGAACCAAGATTTTCTTGATGTTGTTATGGAATAGGGGACAAATCTT 958
QY 911 CAGCCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCTGGCCATGCTTTTGGT 970
DB 959 GAAGGAGAAAGCTCAGTACACCAACCAATCAAGACACTTTTCTGTAATTTGTTTCAAGCT 1018
QY 971 GGAAGTACTCATCAACTGGGACTTTAGAGTGGTCAATTAATTTATTTATGATCACCACCA 1030
DB 1019 GGTACGGACACTTCTTCTTAGTGCATAGAAATGGGACCTTGCAGAAATGATGAAGAACCTT 1078

QY 1031 GAGGTGTTGAGAGGCAAGAGATGATTTGGACACTCAAGTGGGACAGACCCCTTGTGA 1090
Db 1079 GCCATTTTGAAGAAAGACAGCAGAGAAATGGATCAAGTCAATTTGGAAGAAATAGCGCTTTA 1138
QY 1091 AATGAGTCAGACCTTCCAAAGACTTCCATATCTTAGGAAGATCATCTTGGAGACACTTAGG 1150
Db 1139 CTGATCCGATATCCCAATCTCCCTTACTCGAGCAATTTGCAAGAAACATTTCGA 1198
QY 1151 TTGTACCCCGCGCCCAATTTCTAATACCTCATGTGTCTTTCAGAAGATATTAACAATTGAA 1210
Db 1199 AAACACCTTCTACACCAATTAATCTTCTTAGGATCTCGAAGCAACCATGATGCGAT 1258
QY 1211 GGAATCAATATCCACAGACACAAATTTGATCATTAATGTTGGGCGATGACAGAGAT 1270
Db 1259 GGTATTACATACCAAAAACACATAGGCTTAGTGTAAACATATGGGCAATTTGGAAGAGAT 1318
QY 1271 CCTCAGTTGTGGATGATGCCATGCTTTAAACCTGAGAGGT----- 1313
Db 1319 CCCCAAGTTTGGGAAATCCACTAGATTTTAATCCCGAAAGATTTCTTGAGTGGAAAGAAC 1378
QY 1314 -----TTGATGTGGAAGAGAGGAGGAAAAAGTTGGTAGCATTTGGCATGGGAAGAGG 1366
Db 1379 TCCAAGATTGATCTCTCGAGGAGACGATTTTGAATTTGATACCATTTGGTGTGCGACGAGA 1438
QY 1367 GCTTGGCCAGAGAACCCATGGCTATGCAAGAGTGTGAGCTTTACTTTGGGATTTGTTGATT 1426
Db 1439 ATTGTGTCAGGAAACAGAAATGGGAATTTGTAATGTTGGAATATATATTAGGAACCTTTGGTT 1498
QY 1427 CAATGTTTTCAGTGGAAACGAGTAAGTGAGG 1457
Db 1499 CATTCATTTGTTGGAATTTACCAAGTGAAG 1529

RESULT 7

US-09-616-990-1

; Sequence 1, Application US/09616990

; Patent No. 6232109

; GENERAL INFORMATION:

; APPLICANT: KIKUCHI, Yasuhiro

; KIKUCHI, Shigeto

; SHIMADA, Yukihisa

; OHBAVASHI, Masaya

; SHIMADA, Ritsuko

; OKINAKA, Yasushi

; TITLE OF INVENTION: NOVEL PLANT GENES

; NUMBER OF SEQUENCES: 67

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112-3801

; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS Ver3.30

; SOFTWARE: PATENT AID Ver1.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/616,990

; FILING DATE: 14-Jul-2000

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP44963/92

; FILING DATE: 02-MAR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Perry, Lawrence S.

; REGISTRATION NUMBER: 31865

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-218-2100

; TELEFAX: 212-218-2200

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1824 base pairs

;

;

;

;

;

;

;

;

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Petunia hybrida
STRAIN: Falcon Blue
FEATURE:
NAME/KEY: CDS
LOCATION: 116 to 1633
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 1
US-09-616-990-1

Query Match 6.9%; Score 128.6; DB 3; Length 1824;

Best Local Similarity 49.9%; Pred. No. 8e-24;

Matches 435; Conservative 0; Mismatches 399; Indels 37; Gaps 3;

QY 623 ATGATATCGGGAAGAGGTTTACGGAGAGGAGTGAATGAGATGAAGAACTTTCAGGAGCG 582
Db 660 ATATGATCGGACAGAGTGTCTTAAGCAAAAGAGTATTTG-TAGATAAAGGTTGTGAGGTA 718
QY 683 AGGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACTCATGGGCTTGGCTAACAGGGA 742
Db 719 AATGAATTTAAGGACATGTTGTGAGATTATGACATAGCAGGGTATTTTCAACATTGT 778
QY 743 GATCACTTGCCTTCTCAGGTGGTTTCGATTTTTCAGAAATGTCGAGAAGCGCTTAAAGAGT 802
Db 779 GATTTTATCTTCTTGTAGCTTGGATGGATTTTACAAGGATAGAAAAACGAATGAACGT 838
QY 803 ATCAGTAGAGGTACGATTCATCTTGAATAGATCCTTTCATAGAGAACCGTGGCA----- 857
Db 839 TTACATAAAGAGTTTGTGCTTTTATGACAAAGATGTTTGTATGAACAACAAGCAACTACC 898
QY 858 -----GCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACCTGCAAGAGACC 910
Db 899 TATGAACGTAAAGGGAACCAAGATTTTCTGTATGTTGTATGGAATAATGGGCAATTTCT 958
QY 911 CAGCCTCAGTACTACACTGACCAAAATCATCAAAGGCTTGTCTGTGCCATGCTTTTGGT 970
Db 959 GAAGGAGAAAGACTCAGTACAACCAACATCAAAGCACATTTTGTGTAATTTGTTCACAGCT 1018
QY 971 GGAACGTACTCATCACTGGGACTTTAGAGTGGTCTATTATCTTAATTTATTTGATCACCCA 1030
Db 1019 GGTACGGACACTTCTTCTAGTCAATAGAAATGGGCATTTGCAGAAATGATGAAGAACCT 1078
QY 1031 GAGGTGTTGAAGAAGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTGTGA 1090
Db 1079 GCCATTTTGAAGAAAGCAACAGCAGAAATGGATCAAGTCAATTTGGAAGAAATAGCGCTTTA 1138
QY 1091 AATGAGTCAGACCTTCCAAAACCTTCCATATCTTAGGAAGATCATCTTGGAGACACTTAGG 1150
Db 1139 CTCGAATCCGATATCCAAATCTCCCTTACCTCCGAGCAATTTGCAAGAAACATTTTCCA 1198
QY 1151 TTGTACCCCGCGCCCAATTTCTAATACCTCATGTGTCTTCAGAGATATTAACAATTGAA 1210
Db 1199 AAACACCTTCTTACACCAATTAATCTTCTAGGATCTCGAAGCAACCATGATGCGAT 1258
QY 1211 GGAATCAATATCCCGAGACACAATTTGTGATCATTAAATGTTGGGCAATGACAGAGAT 1270
Db 1259 GGTATTACATACCAAAAACACTAGGCTTAGTGTTAATGATATGAGCAATTTGGAAGAT 1318
QY 1271 CCTCAGTTGTGGAATGATGCCATGCTTTTAAACCTGAGAGGT----- 1313
Db 1319 CCCCAAGTTTGGGAAATCCACTAGAGTTTAAATCCCGAAAGATTTCTTGAGTGGAAAGAAC 1378
QY 1314 -----TTGATGTGGAAGAGAGGAGGAAAAAGTTGGTAGCATTTGGCATGGGAAGAGG 1366
Db 1379 TCCAAGATTGATCTCTCGAGGAGACGATTTTGAATTTGATACCATTTGGTGTGCGACGAGA 1438
QY 1367 GCTTGGCCAGAGAACCCATGCTATGCAAGAGTGTGAGCTTTTACTTTGGGATTTGTTGATT 1426
Db 1439 ATTGTGTCAGGAAACAGAAATGGGAATTTGTAATGTTGGAATATATATTAGGAACCTTTGGTT 1498

Qy 1427 CAATGTTTGTACTGGAAACGAGTAAGTGAGG 1457
 ||| | | | | | | | | | | | | |
Db 1499 CATTCA TTGATTGGAAAT TACCAAGTGAAG 1529

RESULT 8

US-08-975-316-17
; Sequence 17, Application US/08975316
; Patent No. 5952486
; GENERAL INFORMATION:
; APPLICANT: BLOKSBERG, Leonard N., HAVUKKALA, Ilkka
; APPLICANT: and GRIERSON, Alastair W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASCEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,316
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,000
; FILING DATE: September 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: SLEATH, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000/1003C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-975-316-17

Db	248	CGTTGGTCTCGGGCGGCTTATTCATGTGCTTCGAATGGGNA	CGAGTTGGCGAGAATTGG	307
Qy	1464	TTGATATGACAGAGAACAAATGGATCACTTGTCAAGGTTAATTCAT	TGGAGGCCATGT	1523
Db	308	TGACACTGTTCGAGGGGACGGGACTCACAATGCCAAGAGAGAGCCAT	TGGAGGCCTTGT	367
Qy	1524	GCAAGGCTCG	1533	
Db	368	GCAAGCGCG	377	

RESULT 9

```

US-09-615-192A-17
; Sequence 17, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
;                Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615.192A
; CURRENT FILING DATE: 2000-07-12
; PRIORITY APPLICATION NUMBER: US 08/975,316
; PRIORITY FILING DATE: 1997-11-21
; PRIORITY APPLICATION NUMBER: US 08/713,000
; PRIORITY FILING DATE: 1996-09-11
; PRIORITY APPLICATION NUMBER: US 09/169,789
; PRIORITY FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-615-192A-17

```

RESULT 10

US-07-912-900-28
; Sequence 28, Application US/07912900

SEQUENCE CHARACTERISTICS:

LENGTH: 1812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-285-309-28

Query Match 6.5%; Score 120.8; DB 1; Length 1812;

Best Local Similarity 50.1%; Pred. No. 8.7e-22;

Matches 437; Conservative 0; Mismatches 397; Indels 38; Gaps 4;

QY 623 ATGATATCGGGAAGAGTTTACCGAGAGGAGAGTGAGATGAAGAAGCTTGAGGAAGCG 682
DB 669 ATATGATCGGACAGTGTCTAAGCAAAAGAGTATTG-TAGATAAAGGTTGTGAGGTA 727
QY 683 AGGAGCTTCAGAGAGACTGTGACAGAAATGTTGGAATCATGGCTTGGCTAACAGGGA 742
DB 728 AATGAATTTAAGACATGTTGTAGAGTTAATGACATAGCAGGTTATTTCACATTGGT 787
QY 743 GATCACTTCCTTCTCAGGTGGTTCGATTTTCAGATGTGGAGAGCGCTTAAAGAGT 802
DB 788 GATTTTATTCCTTTTACCTTGGATGTTTACAGGGATAGAAAAAGAACTGAAACCT 847
QY 803 ATCAGTAAGAGGTACGATTCATCTTGAATAGATCCCTTCATGAGAACCGTGCCTA 857
DB 848 TTACATAAAGAGTTTGTATGCTTTTATGACAAAGATGTTTGTAGAACACAAAGCACTAC 907
QY 858 -----GCAATGACCGGAGATTCATGATCATCTCTCAAACTGCAAGAGACC 910
DB 908 TATGAACGTAAGGGGAAACAGATTTTCTTGATGTTGTATGGAATGCGGACAAATCT 967
QY 911 CAGCCTCAGTACTACACTGACCAAAATCATCAAGGCGCTTCTCGCCATGCTTTTGT 970
DB 968 GAAGGAGAAGACTCAGTACAAACATACAGCACTTTTGTGATTTTGTTCACAGCT 1027
QY 971 GGAATGACTCATCAACTGGGACTTTAGAGTGTCTATATCTAATTTATGATCAACCA 1030
DB 1028 GGTACGAGCACTTCTTCTAGTGAATAGAGTGGCACTTGCAGAAATGATGAGAACCT 1087
QY 1031 GAGGTGTGAAGAGCAAGAGATGAATGGA-CACTCAAGTGGGCAAGACCGCTTGT 1089
DB 1088 GCCATTTTGAAGAAAGCAGAGAAATGGATCAAGGTCAITGGAGAAATAGGCGTTT 1147
QY 1090 AAATGAGTACAGACTTCCAAAATTCATATCTTAGAAGATCATCTTGAGACACTTAG 1149
DB 1148 ACTCGAATCGATATCCCAATCTCCCTTACCTCCGAGCAATTTGCAAGAAACATTCG 1207
QY 1150 GTTGTACCCCGGCCCCCAATTCATACCTCATGTGTCTCAGAGATATTAATTTGA 1209
DB 1208 AAAACACCCCTTACACCAATTAATCTTCTAGGATCTCGAACGAACCATGCTATGCGA 1267
QY 1210 AGGATTCATATCCCGAGACACAAATTTGATCATTAATGTTGGGCGATGAGAGAGA 1269
DB 1268 TGGTTATCATACCAAAAACACTAGGCTTAGTGTATTAATATGGCAATTCGAAGAGA 1327
QY 1270 TCCTCAGTTTGGAAATGATCCCATGCTTTTAAACCTGAGAGT----- 1313
DB 1328 TCCCAAGTTTGGGAAATCCACTAGAGTTTAAATCCCGAAAGATCTTGAGTGAAGAAA 1387
QY 1314 -----TTGATGTGAAGAGGAGGAGAAAAGTTGGTAGCATTTGGCATGGAAGAG 1365
DB 1388 CTCGAAGTTGATCTCGAGGGAACGATTTTGAATTTGATACCAATTTTGGTCTGGAGAG 1447
QY 1366 GGCTTGGCCAGGAGAACCCATGCTATGCAAAAGTGCAGCTTTACTTTGGATTTGTGAT 1425
DB 1448 AATTTGTGAGGAAACAGAAATGGGAATTTGATGGTGGAAATATATATTAGAACTTTGGT 1507
QY 1426 TCAATGTTTGTACTGGAACAGAGTAAGTGGAG 1457
DB 1508 TCATTCATTTGATGGGAAATATACCAAGTGAAG 1539

RESULT 12

US-08-502-046-28
; Sequence 28, Application US/08502046
; Patent No. 5861487

GENERAL INFORMATION:

APPLICANT: Holton, Timothy A.
APPLICANT: Cornish, Edwin C.
APPLICANT: Kovacic, Filipa
APPLICANT: Tanaka, Yoshikazu
APPLICANT: Lester, Diane R.
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
HYDROXYLASE AND USES

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/502,046
APPLICATION NUMBER: US/08/502,046
FILING DATE: 14-JUL-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/285,309

FILING DATE: 03-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8633Z
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1812 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-502-046-28

Query Match 6.5%; Score 120.8; DB 2; Length 1812;

Best Local Similarity 50.1%; Pred. No. 8.7e-22;

Matches 437; Conservative 0; Mismatches 397; Indels 38; Gaps 4;

QY 623 ATGATATCGGGAAGAGTTTACCGAGAGGAGAGTGAGATGAAGAAGCTTGAGGAAGCG 682
DB 669 ATATGATCGGACAGTGTCTAAGCAAAAGAGTATTG-TAGATAAAGGTTGTGAGGTA 727
QY 683 AGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACTCATGGCTTGGCTAACAGGGA 742
DB 728 AATGAATTTAAGGACATGTTGTAGAGTTAATGCAATAGCAGGTTATTTCACATTGGT 787
QY 743 GATCACTTCCTTCTCAGGTGGTTCGATTTTCAGATGTGAGAGCGCTTAAAGAGT 802
DB 788 GATTTTATTCCTTGTGATGTTGATGTTTACAGGATAGAAAAAGAAATGAACCGT 847
QY 803 ATCAGTAAGAGGTACGATTCATCTTGAATAAGATCCTTCATGAGAACCGTGCCTA 857
DB 848 TTACATAAAGAGTTTGTATGCTTTTATGACAAAGATGTTTGTAGAACACAAAGCACTACC 907
QY 858 -----GCAATGACCGGAGAAATTCATGATGATCATCTCTCAAACTGCAAGAGACC 910
DB 908 TATGAACGTAAGGGGAAACAGATTTTCTTGATGTTGTATGGAATGCGGACAAATCT 967

RESULT 14

US-08-285-309-29
 ; Sequence 29, Application US/08285309
 ; Patent No. 5569832
 ; GENERAL INFORMATION:
 ; APPLICANT: Holton, Timothy A.
 ; APPLICANT: Cornish, Edwina C.
 ; APPLICANT: Kovacic, Filipa
 ; APPLICANT: Tanaka, Yoshikazu
 ; APPLICANT: Lester, Diane R.
 ; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
 ; TITLE OF INVENTION: HYDROXYLASE AND USES
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/285,309
 ; FILING DATE: 03-AUG-1994
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DiGiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 86332
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 742-4343
 ; TELEFAX: (516) 742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1755 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-285-309-29

Query Match 6.4%; Score 118.2; DB 1; Length 1755;
 Best Local Similarity 49.0%; Pred. No. 4.1e-21;
 Matches 407; Conservative 0; Mismatches 398; Indels 36; Gaps 2;
 QY 663 TGAAGACGTTGAGGAGCGAGGAGTTCAGAGAGCTGTGACAGAAATGTCGAACTCA 722
 DB 613 TAAATAAGGTGTGAGTAAATGAATTAAGACATGGGTGAGATTAATGACACAG 672
 QY 723 TGGCGTTGGCTAAACAGGAGATCATTGCTTTCCTCAGGTGGTTCGATTTTCAGAAATG 782
 DB 673 CAGGATTTTAAACATTGGTGTATTTTATTCCTTGTAGCTTGGATTTTACAAAGGA 732
 QY 783 TGAAGAGCCGTTAAGAGTATCATGTAAGGTACGATTCATCTTGAATAAGATCTTC 842
 DB 733 TAGAAAAGGAATGAACCGTTTACATAGAAAGTTTGATGCTTTATTGACAAAGATGTTG 792
 QY 843 ATGAGAACCGTCCAGCAATGACCGCAGAAATCCATGATCGATCTCTCTCAAACTGC 902
 DB 793 ATGACACAAGCACTAGTATGAACGTAGGGGAACAGATTTCTTGATTTGTTA 852
 QY 903 AAGAGACCCGCTCAGTACTA-----CACTGACCAATCATCAAAAGGCGCTTG 950
 DB 853 TGGAAAATAGGACAATTTCTGAAGGAGAAAGGCTCAGTACCAACCAATCAAGCACTCT 912
 QY 951 CTCGGCCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGAGTGGTCAATTAT 1010

DB 913 TGCTCAATTTGTTTCAAGCTGGTACAGACACTTCTTCTAGTCAATAGATGGGCACTTG 972
 QY 1011 CTAATTTTATTAATCACCAGAGGTTTGAAGAAGCGCAAGATGAATTTGGACACTCAAG 1070
 DB 973 CAGAGATGATGAAGAACCTTCCCATTTTAAAGAAAGCAAGAGGAATGGATCAAGTCA 1032
 QY 1071 TGGGACAAGACCGCTTTGTTAAATAGTCAAGCTTTCAAAATTCCTCATATCTTAGAAGA 1130
 DB 1033 TTGGAACAATAGGCGTCTGCTCGAATCGGATATCCCAATCTCCCTTACCTCCGAGCAA 1092
 QY 1131 TCATCTTTGAGACACTTAGTTGTACCCCGGCCCCCAATTCCTAATACCTCATGTGCTTT 1190
 DB 1093 TTTGCAAAAGAACATTTGAAAGCACCTTCTACACCAATTAATCTCCCTAGGATCTCGA 1152
 QY 1191 CAGAAGATATTACAAATTGAAGGATTCATATATCCCAAGACACAAATTTGTGATCATTAATG 1250
 DB 1153 ACGAACCATGCAATTCGATGTTTATACATACCAAAAACACTAGGCTTAGTGTAAACA 1212
 QY 1251 GTTGGGCAATGCAAGAGATCTCAGTTGTGGAATGATGCCACATGCTTTTAAACCTGAGA 1310
 DB 1213 TATGGCAATTTGGAAGAGATCCGAAAGTTTGGGAGAACCCACTAGAGTTTTATCTCTGAAA 1272
 QY 1311 GGT-----TTGATGTGGAAGGAGAGGAAAGTTGGTAG 1346
 DB 1273 GGTCTTGTAGTGAAGAAACTCGAAGATTTGCTCGAGGGAAACGCTTTGAAATGATAC 1332
 QY 1347 CATTTGGCATGGGAAGAGGCTTCCCAAGAGAAACCCATGCTATGCAAAAGTGTGAGCT 1406
 DB 1333 CATTTGGTGTGACGAAGATTTGTGAGGAGCAAGAAATCGTAATGGTGAAT 1392
 QY 1407 TTACTTTGGGATTTGATTCATATTTTGTGATGCAAAAGGAGTAAGTAGG 1457
 DB 1393 ATATATTAGGAACCTTTGGTCCATTTGTTGAAATTTACCAAGTGAAG 1443

RESULT 15

US-08-502-046-29
 ; Sequence 29, Application US/08502046
 ; Patent No. 5861487
 ; GENERAL INFORMATION:
 ; APPLICANT: Holton, Timothy A.
 ; APPLICANT: Cornish, Edwina C.
 ; APPLICANT: Kovacic, Filipa
 ; APPLICANT: Tanaka, Yoshikazu
 ; APPLICANT: Lester, Diane R.
 ; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/502,046
 ; FILING DATE: 14-JUL-1995
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/285,309
 ; FILING DATE: 03-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DiGiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 86332
 ; TELECOMMUNICATION INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:29:01 ; Search time 42 Seconds
(without alignments)
1142.576 Million cell updates/sec

Title: US-09-868-546A-2
Perfect score: 2592
Sequence: 1 LIIWVSAYVFLVFLGVKFP.....RLIPLEAMKARPLATKIGI 499

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1287.5	49.7	500	2 C85441	cytochrome P450-11
2	1285	49.6	499	2 T04730	cytochrome P450 ho
3	1241	47.9	497	2 T04731	cytochrome P450 ho
4	1229	47.4	492	2 B85441	cytochrome P450-11
5	1219.5	47.0	495	2 A85441	cytochrome P450-11
6	1212	46.8	543	2 T00513	cytochrome P450 ho
7	1205	46.5	515	2 T00510	probable cytochrom
8	1199.5	46.3	481	2 B96691	probable cytochrom
9	1191.5	46.0	500	2 T52174	cytochrome P450 mo
10	1188.5	45.9	500	2 T04737	cytochrome P450 ho
11	1184.5	45.7	500	2 T52175	cytochrome P450 mo
12	1168.5	45.1	518	2 H85440	cytochrome P450-11
13	1125	43.4	501	2 T04735	cytochrome P450 ho
14	1123	43.3	505	2 T10896	cytochrome P450 (E
15	903.5	34.9	561	2 T49979	cytochrome P450-11
16	888.5	34.3	500	2 T49978	cytochrome P450-11
17	850.5	32.8	513	2 T45624	flavonoid 3'-hydro
18	808.5	31.2	523	2 T04648	cytochrome P450 F1
19	801	30.9	524	2 T04651	cytochrome P450 F1
20	795.5	30.7	543	2 T06533	cytochrome P450 ho
21	784.5	30.3	502	2 T52256	cytochrome P-450LX
22	783	30.2	338	2 T04734	cytochrome P450 ho
23	764.5	29.5	532	2 T05942	cytochrome P450 82
24	762.5	29.4	512	2 T00870	probable cytochrom
25	761.5	29.4	527	2 T07748	probable cytochrom
26	758	29.2	512	2 T00869	probable cytochrom
27	757	29.2	510	2 JC7886	cytochrome P450 92
28	756	29.2	506	2 S38985	flavonoid 3',5'-hy
29	753.5	29.1	498	2 T03634	cytochrome P450 -

RESULT 1

C85441
Cytochrome P450-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001
C:Accession: C85441
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: C85441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-500 <STO>
A:Cross-references: GB:NC_001268; NID:g7270717; PIDN:CAB80400.1; GSPDB:GNO0140
C:Genetics:
A:Gene: At4g37340
A:Map position: 4
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
P:434/Binding site: heme iron (Cys) (axial ligand) #status predicted

ALIGNMENTS

Query Match	49.7%	Score	1287.5	DB 2	Length	500			
Best Local Similarity	49.7%	Pred. No.	1.5e-77						
Matches	247	Conservative	106	Mismatches	129	Indels	15	Gaps	9
Q	9	VLFVFLVGVKVF	---	QSRKLNTPPGPP	-PLPIIGNLNLEQPIHPPFQRMKQYGN-	63			
Q	6	LIFTFLFISLSTFI	IGRIKRPNLPFSP	WALPVIGHURLKPPHRLVLSVSESLGDA	65				
Q	64	-VVSLWFGSRLAV	VISSPTAYOECFTKH	DVALANRLPSLGRKIFYNNNTVGCSCHGHW	122				
Q	66	PIISLRLGNRLV	FVVSSHSLAECECTK	NDVLANRPNLSLAKSHISYCGTIVVTASYGDHW	125				
Q	123	RNLRRITADV	ISTORVHFSFGIRSD	ETKRLMORVLAKNNEEFARVEISMFNDLTY	182				
Q	126	RNLRRIGAVE	IFSALHNSFSIRDR	IEHRLI--ACLSRNS-LEFTKVEKSMFNLTF	182				
Q	183	NNIMRMISCKR	FVGRSEMKNV	EEAREPRTVTEMLMLGLANKGDHLPFLRWP	242				
Q	183	NNIIRMLACK	YCYGDGAE--	DDPEAKRVRELTAEAGCFCGAGNTADYLPILTW--ITGSE	238				
Q	243	KRLKSIKRYD	SILANKILHEN	RASND--RQNSMIDHLKLQETOPOYYTQIIGKLALAML	301				
Q	239	KRIKIASRL	DELFLQGLVDR	REGKEKEQNTWVDHLLCLQETOPEYTTDNIIGIMSLI	298				
Q	302	FGTDSSTGL	BWSNLNHN	HEVLKARDELDTQVQDRLLNESDLPKLPYLKRIILET	361				
Q	299	LACTDTS	AVTLEWTL	SALLNHPQLSKARDEIDNKVGLNRLVEESDLSHLPQLQNVSES	358				
Q	362	LRLYPAPIL	IHPVSS	EDITIEGFNI	PRDTIVINGWGMORPQLWNDATCFKPERFDVE	421			

A:Molecule type: DNA
A:Residues: 1-481 <STO>
A:Cross-references: GB:AE005173; NID:g11054542; PIDN:AAG27841.1; GSPDB:GN00141
C:Genetics:
A:Gene: P28G11.4
A:Map position: 1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:416/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 46.3%; Score 1199.5; DB 2; Length 481;
Best Local Similarity 47.0%; Pred. No. 9.2e-72;
Matches 237; Conservative 101; Mismatches 123; Indels 43; Gaps 11;

QY 2 LVVSYAVLFLVLFLGVKVFQSKRLR---NPPGPP-PLPIGNLNLLEQPIHRRFFORM 57
Db 8 LILITLILTL-----ITLTLKLLKPKPNLPPSPFSLPIIHLRLKLPPLHRLFLFI 62

QY 58 SKOYGN--VVSFWGRLAVVISPTAYQECFTKHDVALANRLPSLSGKYIFYNNTVGS 115
Db 63 SKSLNDSPIFSLRLGNKLVFVSSHSIVSECECTKNDVLANR-P----- 105

QY 116 CSHEHWRNLRRITLADVLSRTQVHSGIRSDTKRLMORLVAKNSNEEFARVEIS 175
Db 106 -----KNLRRIGAVEIFSNHRLNSFYIRDEIRRLIARLSRPNAS-LEFAKVMNS 157

QY 176 MNDLTNNIMRMISGRFYGESEEMKNVEAREFRETVMLEMLGLANKGDLPLFLRW 235
Db 158 MLSLNAFNLIIRVMTGKYYGDGA--DDEAKRVQLIAEAMSCFAGAGAADLPLMLRW 215

QY 236 F-PQNVKRLKISIKRYDSILNKLHNRP-ANDRONSMIDHLKLOETQPOYYTDQII 293
Db 216 ITDF---ERRVKKIAARLDEFFORLVDKRVAKKENTMIDHLLSLQVSQPEYYTDHTI 272

QY 294 KGLALAMFLGTDSTGTLEWSLNLNHPVLKARDELDTQVGDRLNLSDELPLKLY 353
Db 273 KGNLSLILAGTDSAVTLEWALSLLNHPVLSKARDEINQIGDLRLLESIDINLPY 332

QY 354 LRKILLETLYPPAPILIPHVSSDITIEGFNIPRDTIIVINGMGMORDPOLMNDATCF 413
Db 333 LQNVSETLYLPAGPLLVPHISSDCVKGYDMPGCTMLLVNVAIHRDPRLWDPPASF 392

QY 414 KPRFVDEGEKKLVAFGMGRACPEPMAMQSVFTGLLIIQCFDWKRVSEKLDMTEN 473
Db 393 KPERFEKEGTHKLLTGLGRACPGSLARRLVSLSLGSLIQCFEWERIGEVEVDTEG 452

QY 474 NWITLSRLIPLAEMCKARPLATKI 497
Db 453 GGLTMPRAIPLVAMCRARAFVGKI 476

RESULT 9
T52174
cytochrome P450 monooxygenase [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 16-Feb-2001
C:Accession: T52174
R:Mizutani, M.; Ward, E.; Ohta, D.
Plant Mol. Biol. 37, 39-52, 1998
A:Title: Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differ
A:Reference number: 214382; PMID:9620263
A:Accession: T52174
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-500 <MTZ>
A:Cross-references: EMBL:D78606; PIDN:BA28538.1
C:Genetics:
A:Gene: CYP91A1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:301-460/Domain: cytochrome P450 homology <P45>
F:438/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 46.0%; Score 1191.5; DB 2; Length 500;
Best Local Similarity 46.9%; Pred. No. 3.3e-71;
Matches 235; Conservative 100; Mismatches 151; Indels 15; Gaps 7;

QY 4 VVSYAVLFLVLFLGVKVFQSKRLRNIPPPGPP-LPIIGNLNLLEQPIHRRFFORMSKOY- 61
Db 6 VVLFSI-FSLFLIISFKFLPKQNLPPSPGWLPIIHLRLKLPPIHRLRSFSETLD 64

QY 62 ----GNVSVLWFGSRLAVVIS-PTAYQECFTKHDVALANRLPSLSGKYIFYNNTVGS 116
Db 65 HNDGGVMSLRLGSLRVVSSHKVAABECFCGKNDVLANRPQVIIGHGVGYNNAIMAA 124

QY 117 SHGEHWRNLRRITLADVLSRTQVHSGIRSDTKRLMORLVAKNSNEEFARVEISGM 176
Db 125 PYGDHWRNLRLCTIEIFSTRLNCFLVYRTDEVRLISRLSLAGTKY--TVVELKPM 181

QY 177 FNDLTNNIMRMISGRFYGESEEMKNVEAREFRETVMLEMLGLANKGDLPLFLRW 236
Db 182 LMDLTFNNIMRMWTKRYIGEET--TDEEAKRVKLVADVAGANTSGNAVDPILRL- 238

QY 237 DFQNVKRLKISIKRYDSILNKLHNRP-ANDRONSMIDHLKLOETQPOYYTDQIIKGL 296
Db 239 -FSSYENRVKLGEBTDKFLQGLDDKRGQOETGTTMIDHLLVLQKSDIEYYTDQIIKGI 297

QY 297 ALAMFLGTDSTGTLEWSLNLNHPVLKARDELDTQVGDRLNLSDELPLKLYLRK 356
Db 298 ILIMVIAGTNSAVTLEWALSLLNHPVLSKARDEINRVGLDLRIEADLSLPLKXN 357

QY 357 ILETLYPPAPILIPHVSSDITIEGFNIPRDTIIVINGMGMORDPOLMNDATCFKPE 416
Db 358 IVLETLYLPATPLLVPHMASEDCIKSYDMPRGTTLLVNAWAIHRDPTWDDPDSFKPE 417

QY 417 RPDVGEKELVAFGMGRACPEPMAMQSVFTGLLIIQCFDWKRVSEKLDMTENNWI 476
Db 418 RPEKEEAQKLLAFGLGRACPGSLAQRIVGLALGSLIQCFEWERVGVNVEVDMKEGVGN 477

QY 477 TSLRLIPLAEMCKARPLATKI 497
Db 478 TVPKAIPKAUKARPLHKI 498

RESULT 10
T04737
cytochrome P450 homolog F6G17.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 04-Mar-2000
C:Accession: T04737
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15382
A:Accession: T04737
A:Molecule type: DNA
A:Residues: 1-500 <BEV>
A:Experimental source: cultivar Columbia; BAC clone F6G17
C:Genetics:
A:Map position: 4
A:Introns: 295/3
A:Note: F6G17.80
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:297-460/Domain: cytochrome P450 homology <P45>
F:438/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 45.9%; Score 1188.5; DB 2; Length 500;
Best Local Similarity 45.0%; Pred. No. 5.2e-71;
Matches 225; Conservative 120; Mismatches 142; Indels 13; Gaps 9;

QY 7 YAVLFLVLFLGV--KPVFQSKRLRNIPPPGPPPLPIIGNLNLLEQPIHRRFFORMSKOYGNV 64
Db 3 YFTLLPLFLVTSYKFLYSKTRQFNLPDPPGPPFVGHLLHLMKPPPIHRLRYSNOYGP 62

QY 65 VSLWFGSRLAVVISPTAYQECFT-KHDVALANRLPSLSGKYIFYNNTVGSCHGEHWR 123

Db 63 FSLRFGRRVVVITSPSLAESFTGQNDIVLSSPLQLTAKYVAYNNHTTGTAPYGDHWR 122
Qy 124 NLRITLALDVLSTORVHSFSGIRSDETKRLMORLV-LAKNSNE-EFARVEISSMFNDLT 181
Db 123 NLRICQEISSLHRLNFQHIRKDEILRLMTRLSRYTQTSNESNDTHLEPLLSDLT 182
Qy 182 YNNIMRMISGRFYGESEMKVNEAREFRETVTTEMLMGLANKGDLHPLFLRWDFQNV 241
Db 183 FNNIVRMVTKRYGDD--DVNNKEEAEFLFKLVYDIAMYSKANHSADYLPILKLFQ-NKF 239
Qy 242 EKRKLSISKRYDSITLNLKILHENRASNDRONSMIDHLLKLOETQPOYTTDOIIKGLALML 301
Db 240 EKEVKAIGKSMDDILQRLDECRRDKE-GNTVWNHLISLQOQOPEYTTDVIILKGLMWSM 298
Qy 302 FGGTDSSTGTLEWSLSNLLNHPVLEKARDELDTQVQDRLNLSNEDLPKLPYLKILLET 361
Db 299 LAGTETSATVLEWAMANLLRNPEVLEKARSEIDEKIGKDLIDESDIAPVLYQNVVSET 358
Qy 362 LRLYPPAPILPHVSSEDITIEGFNI PRDTIVILINGMQRDPOLMNDATCFKPERF-DV 420
Db 359 FRLFPVAPFLIPRSPDDMKIGGYDVPRTIVMVNAWAIHRDPEIWEPEKFNPDYNDG 418
Qy 421 EGEE--KLVAFGMRACRACPEPMAMQSVSTGLLIIQCFDWKRVSEKLDNTENNWT 477
Db 419 CGSDYYVYKLMFPNGRRCTPCGAGLGQRIVTLALGSLIQCFEWNKGEEMDSSTGLG 478
Qy 478 LSRLLPLEAMCKARPLATKI 497
Db 479 MRKMDPLAMCRPRPINSKL 498

RESULT 11

T52175
cytochrome P450 monooxygenase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 16-Feb-2001
C;Accession: T52175
R;Mizutani, M.; Ward, E.; Ohta, D.
Plant Mol. Biol. 37, 39-52, 1998
A;Title: Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differ
A;Reference number: Z14382; MUID:98281573; PMID:9620263
A;Accession: T52175
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-500 <MIZ>
A;Cross-references: EMBL:D78607; PIDN:BAA28539.1
C;Genetics:

A;Note: CYP91A2
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;297-460/Domain: cytochrome P450 homology <P45>
F;438/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 45.78; Score 1184.5; DB 2; Length 500;
Best Local Similarity 44.8%; Pred. No. 9.6e-71;
Matches 224; Conservative 120; Mismatches 143; Indels 13; Gaps 9;
Qy 7 YAVLFLVLFGV--KVFQSRKLRNIPPGPPPLPIGNLNLLEOPIHRFFORMSKOYGVN 64
Db 3 YFILLPLLVISYKFLYKSTORNLNPPGPPSRPFVGHLMKPPHRLRLQVSNQYGP 62
Qy 65 VSLWFGRLAVISSPTAYQECFT-KHDVALANRLPSLSGKYIFYNNNTTGVSGSHGHW 123
Db 63 FSLRFGRRVVVITSPSLAESFTGQNDIVLSSPLQLTAKYVAYNNHTTGTAPYGDHWR 122
Qy 124 NLRITLALDVLSTORVHSFSGIRSDETKRLMORLV-LAKNSNE-EFARVEISSMFNDLT 181
Db 123 NLRMCQEISSLHRLNFQHIRKDEILRLMTRLSRYTQTSNESNDTHLEPLLSDLT 182
Qy 182 YNNIMRMISGRFYGESEMKVNEAREFRETVTTEMLMGLANKGDLHPLFLRWDFQNV 241
Db 183 FNNIVRMVTKRYGDD--DVNNKEEAEFLFKLVYDIAMYSKANHSADYLPILKLFQ-NKF 239

Qy 242 EKRKLSISKRYDSITLNLKILHENRASNDRONSMIDHLLKLOETQPOYTTDOIIKGLALML 301
Db 240 EKEVKAIGKSMDDILQRLDECRRDKE-GNTVWNHLISLQOQOPEYTTDVIILKGLMWSM 298
Qy 302 FGGTDSSTGTLEWSLSNLLNHPVLEKARDELDTQVQDRLNLSNEDLPKLPYLKILLET 361
Db 299 LAGTETSATVLEWAMANLLRNPEVLEKARSEIDEKIGKDLIDESDIAPVLYQNVVSET 358
Qy 362 LRLYPPAPILPHVSSEDITIEGFNI PRDTIVILINGMQRDPOLMNDATCFKPERF-DV 420
Db 359 FRLFPVAPFLIPRSPDDMKIGGYDVPRTIVMVNAWAIHRDPEIWEPEKFNPDYNDG 418
Qy 421 EGEE--KLVAFGMRACRACPEPMAMQSVSTGLLIIQCFDWKRVSEKLDNTENNWT 477
Db 419 CGSDYYVYKLMFPNGRRCTPCGAGLGQRIVTLALGSLIQCFEWNKGEEMDSSTGLG 478
Qy 478 LSRLLPLEAMCKARPLATKI 497
Db 479 MRKMDPLAMCRPRPINSKL 498

RESULT 12

H85440
cytochrome P450-like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001
C;Accession: H85440
R;anonymous; The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: H85440
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-518 <STO>
A;Cross-references: GB:NC_001268; NID:G7270714; PIDN:CAB80397.1; GSPDB:GN00140
C;Genetics:

A;Map position: 4
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;446/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 45.1%; Score 1168.5; DB 2; Length 518;
Best Local Similarity 46.7%; Pred. No. 1.2e-69;
Matches 239; Conservative 101; Mismatches 147; Indels 25; Gaps 10;
Qy 1 LLVVVSYAVLFLVLFLGVKVFQSRKLRNIPPGPP-PLPIGNLNLLEOP-IHRFFORMS 58
Db 5 LLILATLVAIFTVKIV---LLVTKPNKLNLPSPNICFPPIIHLHLKPLHLRLTSLHS 60
Qy 59 KOYGVNVSILWFGSRLAVISSPTAYQECF-TKHDVALANRLPSLSGKYIFYNNNTTGVSGS 117
Db 61 HSLGPVSLRGLSRLAVISSPTAAEECEFLTKNDIVLANRPFINGKYVAYDTSMVTAP 120
Qy 118 HGEWRNLRITLALDVLSTORVHSFSGIRSDETKRLMORLVLANKSNNEEFARVEISSMF 177
Db 121 YGDWRNLRITLALDVLSTORVHSFSGIRSDETKRLMORLVLANRPFINGKYVAYDTSMVTAP 177
Qy 178 NDLTYNNIMRMISGRFYGESEMKVNEAREFRETVTTEMLMGLANKGDLHPLFLRWFD 237
Db 178 TGLTLNVMRMVMTGKRRFF-EEDDGGKAGISLRELVAEILSAAADNPADELPAURWFD 236
Qy 238 FQNVKRLKLSISKRYDSITLNLKILHENRASNDR--QNSMIDHLLKLOETQPOYTTDOIIK 294
Db 237 YGLVKAIRIGERMDSLLQGLDDEHRANKORLEFNTMIALLLDSQEKEPNYSQDITK 296
Qy 295 GLALAMLFGGTDSSTGTLEWSLSNLLNHPVLEKARDELDTQV--GQDRILNESDLPKL 351
Db 297 GLILAMVVGTDTSALTVEWAMSNLLNHPQILLETTRQNDITQMETSSSRRLKKEEDLVNM 356
Qy 352 PYLRKIILETLRLYPAPILPHVSSEDITIEGFNI PRDTIVILINGMQRDPOLMNDAT 411


```
Db 357 NYLVNVSEILRLYPVAPLPPHVPSSDCVIGFNVPRDITVLVNLWAIHRDSVWDDPT 416
Qy 412 CFKPERFDVEGEK-----KLVAFGMRACGPEPMAMQSVFTLLGLLQCQFDWKRVS 465
Db 417 SFKPERF--EGSDQFHNGKMPFGLGRACPGLSLANRVVGLLGSMLQCFEWSGSG 474
Qy 466 EKLDMTENNWITLSRLLPLEAMCKARPLATKI 497
Db 475 GOVDMTEGPGSLPKAEPLVVTCTREMASEL 506

RESULT 13
T04735
cytochrome P450 homolog F6G17.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jan-2000
C:Accession: T04735
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15382
A:Accession: T04735
A:Molecule type: DNA
A:Residues: 1-501 <BEV>
A:Cross-references: EMBL:AL035601
A:Experimental source: cultivar Columbia; BAC clone F6G17
C:Genetics:
A:Map position: 4
A:Insertions: 292/3
A:Note: F6G17.60
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homolog
F:294-457/Domain: cytochrome P450 homolog <P45>

Query Match 43.4%; Score 1125; DB 2; Length 501;
Best Local Similarity 44.7%; Pred. No. 8.3e-67;
Matches 222; Conservative 120; Mismatches 139; Indels 16; Gaps 10;

Qy 9 VFLVLF-LGVKVFQSKRL-NIPGPP-PLPIIGNLNLLLEQPIHRFFORMSKQGVNV 65
Db 7 ILPLALFLAYKFFFTSKQRYLPSPSYSLPILGHLLAIKPPVHRLFRLNHIPIF 66
Qy 66 SLWFGSLVAVVIGSPYQSCFT-KHDVALANRLPSLSGKYFYNNNTVGCSSHGHRN 124
Db 67 YLGLSRAVVIVSSSLARCFQGVNDVIVSNRPLTSTKIAYNTIATTSYGDHRN 126
Qy 125 LRRITADVLTQVHSGIRSDTKLMQRLVLAQNSNEEFARVEISSMFDNLTYNN 184
Db 127 LRRICSLVSSKRLANFLIRKEEIQRLTSLRSDARVCKE---VELSSILVDLTNN 182
Qy 185 IMRMISGRFYGESEMKNVVEAREFRETVMLEMLGLANKGDHLPFLRFQNVK 244
Db 183 IVRMTVKIYGD--DVSDEEAEFLPKLFTFITNSGARHPGSEYLPFMKIFG-GSPEKE 239
Qy 245 LKSTIKRYDSTLNKLHNENRASNDRQNSMIDHLLKQETOPOYYTDOI IKGLALMLFGG 304
Db 240 VKAAKVIDEMLQRLDECKSDK-DGMTWNHLLSLQDDPEYITDI IKGLMLGIVAS 298
Qy 305 TDSSTGTLEWSNLNHPVYLKKADELDTQVQDRLNLESPLKPLRLKILETLRL 364
Db 299 SETSALTIEWAMASLANHPKVLQVKLEIDELIQDRLIESDIANLPYLVNVVSETLRL 358
Qy 365 YPAPILIPIHVSSEDTIEGNIPTDITVINGMQRDPQLNDATCFKPERPD-VEGE 423
Db 359 HPAAPVLVPRSTAEIDIKIGGYDVRDVTWVMVNAWIHRDPLMTPEPRFNPENFGGE 418
Qy 424 E---KLVAFGMRACGPEPMAMQSVFTLLGLLQCQFDWKRVSSEKLDMTENWITLSR 480
Db 419 KDDVRMLIATGSGRRICPGVGLAHKI VTLALGSLIQCFDWMKKNKEISDMSEGPGMWRM 478
Qy 481 LIPLEAMCKARPLATKI 497
Db 479 MVPLRALCKTRPINKL 495
```

RESULT 14

```
T10896
cytochrome P450 (EC 1.14.-.-) 81B1c - Jerusalem artichoke
C:Species: Helianthus tuberosus (Jerusalem artichoke)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
C:Accession: T10896
R:Cabello-Hurtado, F.; Bataud, Y.; Salaun, J.; Durst, F.; Pinot, F.; Werck-Reichhart,
J. Biol. Chem. 273, 7260-7267, 1998
A:Title: Cloning, expression in yeast and functional characterization of CYP81B1, a p
A:Reference number: Z17204; MUID:98184826; PMID:9516419
A:Accession: T10896
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-505 <CAB>
A:Cross-references: EMBL:AJ000477; NID:g3059128; PIDN:CAA04116.1; PID:g3059129
A:Experimental source: cv. blanc commun
C:Genetics:
A:Gene: CYP81B1c
C:Function:
A:Description: specifically catalyzes the hydroxylation of medium chain saturated fatt
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homolog
C:Keywords: chromoprotein; fatty acid metabolism; heme; iron; metalloprotein; oxidorec
F:301-463/Domain: cytochrome P450 homolog <P45>
F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 43.3%; Score 1123; DB 2; Length 505;
Best Local Similarity 43.5%; Pred. No. 1.1e-66;
Matches 220; Conservative 118; Mismatches 148; Indels 20; Gaps 8;

Qy 2 LVVSVAVLFLVLEGLGVKVFQSKRLNIPGPP-PLPIIGNLNLLLEQPIHRFFORMSKQ 60
Db 6 LUTTTLLLTFTLYLLLR-----RRSSTLPTPIFSLPIIGHLYLKPPLRYRTLAKLSAK 60
Qy 61 YGVNVSLWFGSLVAVVIGSPYQSCFTKHDVALANRLPSLSGKYFYNNNTVGCSSHGE 120
Db 61 HGQILRLQLGPRRLVIVSSPSAAEECTKNDIVFANRPKMLFGKIIGVNVTSLSWSPYGD 120
Qy 121 HVRNLRRITADVLTQVHSGIRSDTKLMQRLVLAQNSNEEFARVEISSMFDNL 180
Db 121 NWRNLRIASIELSTHRLNEFHDIRVEETRLIIQKLSACNSGS---SQVTWKFSFYEL 177
Qy 181 TYNMIRMISGRFY-GESEMKNVVEAREFRETVMLEMLGLANKGDHLPFLRFDFQ 239
Db 178 TLNVMRMISGRFYFGDNPDEL--EEGRFMDLDETFLAGASNVGYLPLVSLGLGVK 235
Qy 240 NVEKRLKSTIKRYDSTLNKLHNENRASN-----DRQNSMIDHLLKQETOPOYYTDOI IK 294
Db 236 GLEKKILKLEKRDVFFQGLIDQLRKSKGTEDVNVKKTMIELLLSLOETEPEYTTDAMIR 295
Qy 295 GLALMLFGTDSSTGTLEWSNLNHPVYLKKADELDTQVQDRLNLESPLKPLPYL 354
Db 296 SFVLVLLAAGSDTSAGTMEWNVMLNHPQVLKKAQNEIDSIVGKNCLVDESIPNLPYL 355
Qy 355 RKILLETLLRYPAPILIPHVSSDTIEGNIPTDITVINGMQRDPQLNDATCFK 414
Db 356 RCILNETLLYPAGPLVLPVPEASSCCVGVGVNPRGTILIVNQWAIHHDPKVWDEPETFK 415
Qy 415 PERFD-VEGEK--KLVAFGMRACGPEPMAMQSVFTLLGLLQCQFDWKRVSSEKLDMT 471
Db 416 PERFEGLETRDQGLKLLPFGSGRRSCPGEGSLAVRLMGWTLGSLIIQCFDWEETSEELVDMT 475
Qy 472 ENNWITLSRLIPLEAMCKARPLATKI 497
Db 476 EGPGLTMPKRAIPLVAKCKPRVENTNL 501

RESULT 15
T49979
cytochrome P450-like protein - Arabidopsis thaliana
N:Alternate names: protein F12B17.50
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 16-Feb-2001
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:21:43 ; Search time 25 Seconds
(without alignments)

938.653 Million cell updates/sec

Title: US-09-868-546A-2

Perfect score: 2592

Sequence: 1 LILVVSYAVLFLVFLGVKFP.....RLIPLEAMCKARPLATKIGI 499

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1655.5	63.9	499	1 C81E_GLYEC	P93147 glycyrrhiza
2	1196.5	46.2	502	1 C81D_ARATH	O950k2 arabidopsis
3	1188.5	45.9	500	1 C81F_ARATH	O65790 arabidopsis
4	850.5	32.8	513	1 F3PH_ARATH	O950k2 arabidopsis
5	817	31.5	544	1 C821_PEA	O43068 pisum sativ
6	762.5	29.4	512	1 C7C2_ARATH	O64637 arabidopsis
7	761.5	29.4	527	1 C823_SOYBN	O49858 glycine max
8	758	29.2	512	1 C7C1_ARATH	O64636 arabidopsis
9	756	29.2	506	1 C751_PETHY	P48418 petunia hyb
10	753	29.1	471	1 C771_PEARAE	P24485 persea amer
11	746.5	28.8	512	1 F3PH_PETHY	O950k2 arabidopsis
12	738.5	28.5	508	1 C753_PETHY	P48419 petunia hyb
13	738	28.5	510	1 C933_SOYBN	O81973 glycine max
14	735	28.4	522	1 C822_SOYBN	O81972 glycine max
15	731.5	28.2	511	1 C716_NEPPA	O04164 nepeta race
16	730	28.2	523	1 C756_CAME	O04773 campanula m
17	729	28.1	510	1 C7DA_SOYBN	O48923 glycine max
18	727	28.0	513	1 C752_SOLME	P37120 solanum mel
19	719.5	27.8	511	1 C7C4_ARATH	O64635 arabidopsis
20	715.5	27.6	525	1 C824_SOYBN	O49859 glycine max
21	713	27.5	490	1 C76B_HELTU	O23976 helianthus
22	709	27.4	504	1 C724_ARATH	O65786 arabidopsis
23	707.5	27.3	512	1 C981_SORBI	O48956 sorghum bic
24	707	27.3	502	1 C932_SOYBN	O42799 glycine max
25	704.5	27.2	509	1 C931_SOYBN	O42798 glycine max
26	704	27.2	508	1 C983_ARATH	O22203 arabidopsis
27	702.5	27.1	497	1 C71C_ARATH	O49340 arabidopsis
28	701.5	27.1	499	1 C719_SOYBN	O81970 glycine max
29	693.5	26.8	490	1 C71P_ARATH	O950k2 arabidopsis
30	687	26.5	516	1 C754_GENTR	O96591 gentiana tr
31	685.5	26.4	502	1 C718_MENPI	O42716 mencha pipe
32	684.5	26.4	509	1 C822_SOYBN	O48922 glycine max
33	681.5	26.3	510	1 C755_EUSGR	O96418 eustoma gra

34	678	26.2	490	1 C71M_ARATH	O950k2 arabidopsis
35	677.5	26.1	495	1 C71K_ARATH	P93149 glycyrrhiza
36	676	26.1	523	1 C981_GLYEC	O4790 eustoma gra
37	673.5	26.0	510	1 C757_EUSGR	O49340 arabidopsis
38	672.5	25.9	497	1 C71D_ARATH	O950k2 arabidopsis
39	671	25.9	490	1 C71J_ARATH	O950k2 arabidopsis
40	669.5	25.8	501	1 C72P_ARATH	O950k2 arabidopsis
41	669.5	25.8	505	1 C762_SOLME	P37122 solanum mel
42	669	25.8	505	1 C712_SOLME	P37118 solanum mel
43	668.5	25.8	483	1 C71N_ARATH	O950k2 arabidopsis
44	666	25.7	504	1 C7D8_SOYBN	O81974 glycine max
45	664.5	25.6	486	1 C71O_ARATH	O950k2 arabidopsis

ALIGNMENTS

RESULT 1

CB1E_GLYEC STANDARD; PRT; 499 AA.
 ID C81E_GLYEC
 AC P93147;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 81E1 (EC 1.14.-.-) (Isoflavone 2'-hydroxylase) (P450
 91A4) (CYP 8E-3).
 DE CYP81E1 OR CYP91A4.
 GN Glycyrrhiza echinata (Licorice).
 OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrrhiza.
 OX NCBI_TaxID:46348;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Akashi T., Aoki T., Kameya N., Nakamura I., Ayabe S.-I.;
 RT "Two new cytochrome P450 cDNAs from elicitor-induced Licorice
 (Glycyrrhiza echinata L.) cells".
 RL (in) Plant Gene Register PR97-167.
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=99009223; PubMed=9790908;
 RA Akashi T., Aoki T., Ayabe S.-I.;
 RT "CYP81E1, a cytochrome P450 cDNA of licorice (Glycyrrhiza echinata
 L.), encodes isoflavone 2'-hydroxylase".
 RL Biochem. Biophys. Res. Commun. 251:67-70(1998).
 CC -!- FUNCTION: CATALYZES THE HYDROXYLATION OF ISOFLAVONES, DAIDZEIN AND
 FORMONONETIN, TO YIELD 2'-HYDROXYISOFLAVONES, 2'-HYDROXYDAIDZEIN,
 AND 2'-HYDROXYFORMONONETIN, RESPECTIVELY.
 CC -!- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF ISOFLAVONOID-DERIVED
 ANTIMICROBIAL COMPOUNDS OF LEGUMES.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB001379; BAA22422.1; --
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; p450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 DR Oxidoreductase; Monooxygenase; Membrane; Heme.
 FT METAL 436 436
 SQ SEQUENCE 499 AA; 57191 MW; 4D03AFAC818C04DB CRC64;

Query Match 63.9%; Score 1655.5; DB 1; Length 499;
 Best Local Similarity 60.6%; Pred. No. 5.5e-88;
 Matches 301; Conservative 100; Mismatches 93; Indels 3; Gaps 1;

QY 1 LLVVSVAVLVLFLGVKVFQSRKLNRTPPGPPPLPIIGNLLEQPIHRRFFQMSKQ 60
 Db 3 ILSLSYSYFLALFTFNFIIVIRKFNKPLPPGPPSLPIIGNLHLKRLPHRTFKGLSEK 62
 QY 61 YGNVSLWFGSRLAVVSSPTAYQECFTKHDVALANRLPSLSKGYFYNNYTTVGSCHGE 120
 Db 63 YGHVSLWFGSRLVWVSSASEFCQCFKTDVVLNANRPFSLSKGYFYNNYTTVGSCHGE 122
 QY 121 HWNLRRITDALDVLSTQVHSGIRSDTKRLMORVLVLAKNSEEFARVEISSMFDNL 180
 Db 123 HWNLRRITDALDVLNHRINSFSGIRSDTKRLMORVLVLAKNSEEFARVEISSMFDNL 179
 QY 181 TYNNIMRMISCKPFGYSEEMKNVVEARBPRTVTWMLMGLMGLKGDHLPFLRFDQFN 240
 Db 180 TFNNIMRMISCKPFGYSEEMKNVVEARBPRTVTWMLMGLMGLKGDHLPFLRFDQFN 239
 QY 241 VEKRLSKSKYDSILNKLILHNANRSDRONSMIDHLLKLQETQPOYDYOITKGLALAM 300
 Db 240 LEKRLKDIKSTDAFLRGLLEHRTKERRANTMIDHLLNLDQSOPEYITDQILKGLALAM 299
 QY 301 LFGGTSSTGTLWLSNLSNLNHPVLKARDELDTQVQDRLNLSDDLKPLPYLRKIILE 360
 Db 300 LLAGTSSAVTLEWMSNLSNLNHPVLKARDELDTQVQDRLNLSDDLKPLPYLRKIILE 359
 QY 361 TRLYPAPLILPHVSEDTITGPNIPRTIVINGWQORPOLWMDATCPKPRFDV 420
 Db 360 TRLYPAPLILPHVSEDTITGPNIPRTIVINGWQORPOLWMDATCPKPRFDV 419
 QY 421 EGREKLVAGMGRACPGPMAMOSVFTGLLLOCFDWKRYSEKLDNTNNWTLR 480
 Db 420 KGELEKLVAGMGRACPGPMAMOSVFTGLLLOCFDWKRYSEKLDNTNNWTLR 479
 QY 481 LIPLEAMCKARPLATKI 497
 Db 480 LVLPLKAMCKSRPVINKV 496

RESULT 2
 C81D ARATH
 ID C81D ARATH STANDARD; PRT; 502 AA.
 AC Q9FG65; O65789;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 81D1 (EC 1.14.-.-).
 GN CYP81D1 OR CYP91A1 OR AT5G36220 OR T30G6.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asanizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yanada M., Yasuda M., Sato S., de la Baetide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shan R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Etian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,

RA Rampsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
 RA Weizenecker T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villaroel R., Gielen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rüd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransch P.F.,
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:823-826(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 SSP consortium (Salk/Stanford/PGECC).";
 RN Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 3-502 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98281573; PubMed=9620263;
 RA Mizutani M., Ward E., Ohta D.;
 RT "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of
 cDNAs, differential expression, and RFLP mapping of multiple
 cytochromes P450.";
 RT Plant Mol. Biol. 37:39-52(1998).
 RL
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB026661; BAB09361.1; --
 CC EMBL; D78606; BAA28538.1; --
 CC EMBL; AY093766; AAM10388.1; --
 CC HSSP; F14779; 1JP2.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Transmembrane; Heme; Multigene family.
 FT TRANSMEM 6 26
 FT METAL 440 440 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 121 121 T -> A (IN REF. 2).
 SQ SEQUENCE 502 AA; 56723 MW; 1A348A8DD76A41B0 CRC64;
 Query Match 46.2%; Score 1196.5; DB 1; Length 502;
 Best Local Similarity 47.1%; Pred. No. 1.1e-61;
 Matches 236; Conservative 100; Mismatches 150; Indels 15; Gaps 7;
 QY 4 VYVSAYVAVLVLFLGVKVFQSRKLNIPGPPPLPIIGNLLEQPIHRRFFQMSKQY- 61
 Db 8 VVLYSI-FSLIIFLIISFKPLKPKQNLPPSPGWLPIIGHRLKLPPIHRTURSFETLD 66
 QY 62 ----GNVSLWFGSRLAVVSS-PTAYQECFTKHDVALANRLPSLSKGYFYNNYTTVGS 116
 Db 67 HNDGGVWMSRLGSLVYVSSSHKVAABECFKNDVLANRPQVILGHVGYNNYNTWIAA 126
 QY 117 SHGEHWRNLRRTALDVLSTQVHSGIRSDTKRLMORVLVLAKNSEEFARVEISSM 176
 Db 127 PYGDHWRNLRRLCTIEIFSTRNLNCFVYRTDEVRLISRLSLAGTKK---TVVELKPM 183
 QY 177 FNDLTNNIMRMISGKRFYGESEEMKNVVEARBPRTVTWMLMGLMGLKGDHLPFLRW 236
 Db 184 LMDLTFNNIMRMWTKRYGGEET--TDEEAKRVKLVADVAGNANTSSGNNAVYVPLRL- 240
 QY 237 DFQNVKRLKLSIKRYSIDSLINKILHNANRSDRONSMIDHLLKLQETQPOYDYOITKGL 296
 Db 241 -FSSYENRVKLGEEETDKFLOGLIDDKRGQEQETGTTMIDHLLVLQKSDIEYVTDQIKGI 299

QY 297 ALAMLEGGTDSSTGCTLEMSLNILNHPVILKARDELDTQVGDRLNLSLDLPKLYLRK 356
 DB 300 ILIMVAGTNTSNTTLEWALSNILNHPDVISKARDEIDNRVGLDRIEADLSLELYLRK 359
 QY 357 IILETLRLVPPAPILIPHYSSSEDTTEGENIPRDTTVINGMGQMDPOLWNDATCFKPE 416
 DB 360 IVLETLRLHPATLIPVPHMASEDCKTGSYDMPRGITLLVNAWAIHDPNTWDDPSFKPE 419
 QY 417 RFDVEGEKELVAFMGRRCPGEPMAQSVSTLGLLIQCOPKRVSEBKLDMTENNMI 476
 DB 420 RPEKEEAKLAFGLRRACPGSLAQIRVGLGSLGSLIOCFEWRVGVNDMEKGVGN 479
 QY 477 TSLRLIPLNEMCKARPLATKI 497
 DB 480 TVPKAIPLKAICARPLFKHI 500

RESULT 3
 ID CBIF ARATH STANDARD; PRT; 500 AA.
 AC 065790; Q9S2U3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 81P1 (EC 1.14.-.-)
 GN CYP81P1 OR CYP81A2 OR AT4G37430 OR P6G17.80.
 OS Arabidopsis thaliana (Mouse ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Seedling;
 RX MEDLINE=98281573; PubMed=9620263;
 RA Mizutani M., Ward E., Ohta D.,
 RA "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of
 RT cDNAs, differential expression, and RFLP mapping of multiple
 RT cytochromes P450."
 RL Plant Mol. Biol. 37:39-52(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083486; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Welljens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Mociaman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernselman S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann T.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dese C., de Haan M., Maare A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Hezrl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Argüel F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller K., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedror F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,
 RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,

RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stokeling T., Kalicki J., Graves T., Abbott A., Scott K., Johnson D.,
 RA Latreille P., Courtney L., Cloud J., Miller N., Greco T., Kemp K.,
 RA Minx P., Bentley D., Fulton B., Weller N., Pepin K., Hillier L.,
 RA Krammer J., Fulton L., Mardis E., Dante M., Geisel C., Layman D.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonio B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:769-777(1999).
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D78607; BAA28539.1; -.
 DR EMBL; AL035601; CAB38210.1; -.
 DR PIR; T04737; T04737.
 DR PIR; T52175; T52175.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Membrane; Heme.
 FT METAL 438 438 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 106 106 A -> V (IN REF. 1).
 FT CONFLICT 127 127 I -> M (IN REF. 1).
 FT CONFLICT 140 140 N -> I (IN REF. 1).
 FT CONFLICT 454 454 S -> T (IN REF. 1).
 SQ SEQUENCE 500 AA; 57555 MW; 0FB453D2070EA2EA CRC64;
 Query Match 45.9%; Score 1188.5; DB 1; Length 500;
 Best Local Similarity 45.0%; Pred. No. 3e-61;
 Matches 225; Conservative 120; Mismatches 142; Indels 13; Gaps 9;
 QY 7 YAVLFVLVLGV--KPVFQSRKLRNIPGPPPLPIIIGNLNLEQPIHRPFQRMKQYGVN 64
 DB 3 YFILLPLFLVLSYKFLSKTQRFNLPGPPSPFPVGHLLMKPPIHRLRLQYSNQYGP 62
 QY 65 VSLWFGSLRAVVISPTAYQECFT-KHVDALANRLPSLSGKYIFYNNTVTGSCSGHGHWR 123
 DB 63 FSLRFGSRVWVITPSLAQESFTGQNDIVLSSRPLQLTAKYVAYNHTTGTAPYGDHWR 122
 QY 124 NLRRTALDVLSTQVHSPSGIRSDTEKLMQRLV-LAKNSNE-BEFARVELSSMENDLT 181
 DB 123 NLRRCISQBELSHRLINFQHIRKDEILRMLRLSRYTQTSNESNDFTHIEPLSLDLT 182
 QY 182 YNNIMRMISGKIFYGESEMKVKEAREPREVTVMELMLGLANKGDHLPFLRWDFQNV 241
 DB 183 FNNIVMVTGKRYGD--DVNNKEAELEFKLVYDIAMYSNGANHSADYLPILKLFQ-NKF 239
 QY 242 EKRLKSISKRYDSILNKLILHENRASNDRQNSMIDHLKLOEQTPQYVTOIITKGLALML 301
 DB 240 EKEVKAIGKSMDDILQLRLDECKRDKE-GNTVWVNLISLQOQPEYTYTDVIKGLMWSNM 298
 QY 302 FCGTDSSTGCTLEMSLNILNHPVILKARDELDTQVGDRLNLSLDLPKLYLRKILET 361
 DB 299 LAGTETSATVLEWAMANLRLNPEVLEKSEIDEKIGDKRLIDESIAVLQNVVSET 358
 QY 362 LRLYPPAPILIPHYSSSEDTTEGENIPRDTTVINGMGQMDPOLWNDATCFKPERF-DV 420

Db 359 FRLFPVAFPLIPRPTDMDKIGGYDVPDRDTVMVNAWAIHRDPEIWEPEKFPNDRYNDG 418
 QY 421 EGEE--KLVAFQMGRRACGPEMAMQSVSFTLGLLIQCQFDMKRVSEKLDNTENNWIIT 477
 Db 419 CGSDYYVYKLPFGNGRTCPGAGLGQRIVTILGSLIQCQFENVKGEEMDSSTGJG 478
 QY 478 LSRLLPLEAMCKARPLATKI 497
 Db 479 MRKMDPLRAMCRPRPINSKL 498

RESULT 4
 F3PH ARATH
 ID F3PH ARATH STANDARD; PRT; 513 AA.
 AC Q9SD85;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Flavonoid 3'-monooxygenase (EC 1.14.13.21) (Flavonoid 3'-hydroxylase)
 DE (AtF3'H) (Cytochrome P450 75B1) (TRANSPARENT TESTA 7 protein).
 GN CYP75B1 OR F3'H OR TT7 OR AT5G07990 OR F13G24.190.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia, and cv. Landsberg erecta;
 RX MEDLINE=20483349; PubMed=11030433;
 RA Schoenbohm C., Martens S., Eder C., Forkmann G., Weishaar B.;
 RT "Identification of the Arabidopsis thaliana Flavonoid 3'-hydroxylase
 RT gene and functional expression of the encoded P450 enzyme.";
 RL Biol. Chem. 381:749-753(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RX Cordier T.D., Barri-Rewell G., Brugliera F., Cobbett C., Holton T.A.;
 RA "Isolation of a flavonoid 3'-hydroxylase gene corresponding to the Tc7
 RA locus of Arabidopsis thaliana.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RX Saslowsky D., Winkler-Shirley B.;
 RA "Sequence of flavonoid 3'-hydroxylase (F3'H).";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfling T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney J., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowmatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
 RA Rampeger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
 RA Weitzenecker T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel J., Gielen J., Ardiles W.,

RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:823-826 (2000).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=21382915; PubMed=11489181;
 RA Saslowsky D., Winkler-Shirley B.;
 RT "Localization of flavonoid enzymes in Arabidopsis roots.";
 RL Plant J. 27:37-48(2001).
 CC -!- FUNCTION: Catalyzes the 3'-hydroxylation of the flavonoid B-ring
 CC to the 3',4'-hydroxylated state. Convert naringenin to eriodictyol
 CC and dihydrokaempferol to dihydroquercetin.
 CC -!- CATALYTIC ACTIVITY: A flavonoid + NADPH + O(2) = 3'-
 CC hydroxyflavonoid + NADP(+) + H(2)O.
 CC -!- PATHWAY: Flavonoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: High expression in siliques and to a lower
 CC extent in stems, flowers and senescing leaves.
 CC -!- INDUCTION: By UV light treatment.
 CC -!- MISCELLANEOUS: May act as a membrane anchor for localization of
 CC other, soluble, flavonoid enzymes to the endoplasmic reticulum.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF271651; AAG16746.1; -
 CC EMBL; AF271650; AAG16745.1; -
 CC EMBL; AF155171; AAF73253.1; -
 CC EMBL; AF241646; AAF60189.1; -
 CC EMBL; AF241643; AAF60189.1; JOINED.
 CC EMBL; AF241644; AAF60189.1; JOINED.
 CC EMBL; AF241645; AAF60189.1; JOINED.
 CC EMBL; AL133421; CAB62611.1; -
 CC PIR; T45624; T45624.
 CC InterPro; IPR001128; Cytochrome_P450.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR00385; P450.
 CC PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC Flavonoid biosynthesis; Oxidoreductase; Monooxygenase;
 CC Endoplasmic reticulum; Heme; NADP; Signal-anchor; Transmembrane.
 CC TRANSMEM 1 21
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC (POTENTIAL).
 CC DOMAIN 22 513
 CC CYTOPLASMIC (POTENTIAL).
 CC METAL 445 445
 CC IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC SEQUENCE 513 AA; 56786 MW; C0C740FBE4559C40 CRC64;
 CC
 CC Query Match 32.8%; Score 850.5; DB 1; Length 513;
 CC Best Local Similarity 35.5%; Pred. No. 7e-42;
 CC Matches 182; Conservative 100; Mismatches 193; Indels 37; Gaps 8;
 CC
 CC QY 2 LVVSVAVFLVFLVFLGKVFQSKLRN---IPGPGPPPLPIIGNLLEQPIHRRFQRM 57
 CC Db 6 LTILLATVFLIL-----RIFSHRRNRSHNNRUPPGNPWPIIGNLPHMGTKPHRTLSAM 60
 CC QY 58 SKOYGNVSVLWFGSLRVLVISPTAYQECFTKHDVALANRLPSLSGKYIFPNNTTVGSCS 117
 CC Db 61 VTTYGPIILHLRLGFDVVVAASKVAEQFLKIHDFANFASPPNPGKAMAYNQDLVFP 120
 CC QY 118 HGBHWNRLARITADVLSQTVHSFGIRSDFTKRLMQRVLVLAKNNEEFARVEISSMF 177
 CC Db 121 YGHRWRLLRKISSVHLFSKALEDFKVRQEEVGTTLRELVRVGTK-----PVLGLQLV 174
 CC QY 178 NDLTYNNIMRMISGKFPYGESESEKMKVEAREPREVTMLELMLGLANKGDHLPELRWFD 237
 CC Db 175 NMCVVALGEMIGRRRLFGADADHK----ADEFRSMVTMMALAGVFNIGDFVFSLOWLD 230

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AC003680; AAC06158.1; -;
 DR PIR; T00870; T00870.
 DR HSP; P00179; LDT6.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR Oxidoreductase; Monooxygenase; Heme; Transmembrane; Multigene family.
 FT TRANSMEM 3 23 POTENTIAL.
 FT METAL 451 451 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL 451 451 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 512 AA; 57256 MW; 4PF5ALBEIC24C798 CRC64;

Query Match 29.4%; Score 762.5; DB 1; Length 512;
 Best Local Similarity 34.2%; Pred. No. 7.6e-37;
 Matches 175; Conservative 110; Mismatches 187; Indels 39; Gaps 12;
 1 LLVVVAVLFLVLFLGVKVFQSRKLRNIPPGPPPLPIIGNLLSQPIHRFFQRMKQ 60
 9 LFLPFCFLSPFIIFITTRPSRKRKVVSPGPPRLPIIGNIHLVGRNPHHSFADLSKT 68
 61 YGVVSLWFGSLAVISSPAYOECTKHVALANRLPSLSGKYIFYNNTV---GSC 116
 69 YGPIMSLKFSGSLNTVVVTSPEAREVLTQDILSSRTPTNSIRSNHDKVSVVWLPSS 128
 117 SHGHHWRLRITDALDVLSTQRVHSGIRDETKRLMQRLVLAKNSEEEFARVEISSM 176
 129 S---RWLLRLKLSATQLSPQRIEATKTLRENKVELVS--FMSESSEREE--AVDISR- 180
 177 FNDITYNNIMISGKRYGESEKMNVEARE--FRETVMELMGLANKGDHLPFLR 234
 181 ---ATFITALNIISLNF---SVDLGNYSNKGVFQDTVIGVMEAVGNPDAAFFPFLG 234
 235 WFDQNVKRLKLSIKRVDLSIL-----NKILHENRASNDRONSMIDHLLKQI- 284
 235 FLDLQGNKRLKACSERLKFVFRGFDIAKAEKSLRDTNSKQVREDFVDVLLDTGDE 294
 285 PQYTDQIIKGLALAMLFPGTDSSTGTLEWSLSNLLHNPVLEKARDELDTQVQDRLN 344
 295 AELNTDIVH-LLLDLFCAGTDTNSSTVEWMAELLRNPTWKAQAEIDCVIGQGVVE 353
 345 ESDLPKLPYLKRIILETLRLYPPAPILPHVSSDITIEGNIIPRDTIVINGMQRDP 404
 354 ESDISALPYLOAVVKETFLHPAAPLLVPRKAESDVEVLGFWVPKDTQVFVNVWAIGRP 413
 405 QLWNDATCFKPERE-----DVEGEKKLVAFGMGRACPGPEPMQSVSFTLGLLIQCFD 459
 414 NVWNSRFRKPERFLGKDIDLRGRDYELTPFGAGRRICPGPLAVKTVPLMLASLLYSFD 473
 460 WK---RVSEEEKLDMTENNWTLSRLIPLEAM 487
 474 WKLPNGVSGEDLMDMTETGLLHKTNPPLHAV 504

RESULT 7

C823 SOYBN STANDARD; PRT; 527 AA.
 ID C823 SOYBN STANDARD; PRT; 527 AA.
 AC O49858;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 82A3 (EC 1.14.-.-) (P450 CP6).
 GN CyP82A3.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Harosoy 63;

RX MEDLINE=98311068; PubMed=9648734;
 RA Schopfer C.R., Ebel J.;
 RT "Identification of elicitor-induced cytochrome P450s of soybean
 (Glycine max L.) using differential display of mRNA.";
 RL Mol. Gen. Genet. 258:315-322(1998).
 CC -|- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; Y10982; CAA71876.1; -;
 DR PIR; T07748; T07748.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR Oxidoreductase; Monooxygenase; Membrane; Heme.
 FT METAL 464 464 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 527 AA; 59822 MW; CA3429E87B202210 CRC64;

Query Match 29.4%; Score 761.5; DB 1; Length 527;
 Best Local Similarity 34.6%; Pred. No. 9e-37;
 Matches 175; Conservative 103; Mismatches 199; Indels 29; Gaps 7;
 8 AVLFLVLGVKVFQSRKLRNIPPGPPPLPIIGNLL--EOPHRRFFQRMKQGVNV 65
 20 SLIFLCFLYRK-----NSRGKQAPVVGAWPIGLHLSLNGSQTPHKVLCALADKTGPLF 75
 66 SLWFGSLAVISSPAYOECTKHVALANRLPSLSGKYIFYNNTVSGSCHEWRNL 125
 76 TIKLGMKPALVLSNWEWSKELETTNDLAVSSRPKLAVAVMSVNOAFVGLAPYGYREL 135
 126 RRTALDVLSTORVHSGIRDETKRLMORLV---LAKNSNEEFARVEISSMFDLTY 182
 136 RKIVTPEFLSNRIEQRNHRVSEVTSIKELFDIWSNGNKNESRYTLVDIKQMLAYLTF 195
 183 NNIMRMISGKRYGESEKMNVEAREFRETVMELMGLANKGDHLPFLRWFDFQNV 242
 196 NMVVMVGRKRYFG-VMHVEGKDKAORFMKIRFNNLMGTFTVADGVPCLRMLDGGHE 254
 243 KRLKLSIKRYDSTILNKLHNHNR-----ASNDRONSMIDHLLKLOSTQPYTDDI 292
 255 KAMNAKEVDKLLSEWSEHQRKLLGENVESDRDFMDVMISALNGAQ--ICAFDADTI 312
 293 IKGLALAMLFPGTDSSTGTLEWSLSNLLHNPVLEKARDELDTQVQDRLNLSDLPLKLP 352
 313 CKATSEILIGGTDSTAVTLTWALSLLRNPLALGAKAEIDMOIGKDEYIRESDISKLV 372
 353 YLRKIILETLRLYPPAPILPHVSSDITIEGNIIPRDTIVINGMQRDPOLWNDATC 412
 373 YLOAIVKETLRLYPPAPFPSPFPTEFTNCILGGVHIKKKTELHNLKIHDPSPVMSDPLE 432
 413 FKPERF-----DVEGEKKLVAFGMGRACPGPEPMQSVSFTLGLLIQCFDWRKVS 465
 433 FKPERFLTHKDVLDLGRHNFELLPFGSGRRVCAGMSGLNMMVHFTLANLHLSFDILNPSA 492
 466 EKLDMTENNWTLSRLIPLEAMCKAR 491
 493 EPVDMTEFFGFTNTKATPLEILVKPR 518

RESULT 8
 C7C1 ARATH STANDARD; PRT; 512 AA.
 ID C7C1 ARATH STANDARD; PRT; 512 AA.
 AC O64636; O65783;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

Cytochrome P450 76C1 (EC 1.14.-.-).
 CYP76C1 OR AT2G45560 OR F17K2.9.
 Arabidopsis thaliana (Mouse-ear cress).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 [1]
 SEQUENCE FROM N.A. (ISOFORM LONG).
 STRAIN=cv. Columbia; TISSUE=Seedling;
 MEDLINE=98281573; PubMed=9620263;
 Mizutani M., Ward E., Ohta D.;
 "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of
 cDNAs, differential expression, and RFLP mapping of multiple
 cytochromes P450";
 RT
 Plant Mol. Biol. 37:39-52(1998).
 [2]
 SEQUENCE FROM N.A.
 STRAIN=cv. Columbia;
 MEDLINE=20083487; PubMed=10617197;
 Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 Buell C.R., Ketchum K.A., Lee J.J., Rinning C.M., Koo H.L.,
 Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
 Tallon L.J., Gali J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 Goodman W.C., Somerville C.R., Copenhaver G.P., Preuss D.,
 Nierman H.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 Venter J.C.;
 "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 RT
 Nature 402:761-768(1999).
 [3]
 SEQUENCE FROM N.A. (ISOFORM SHORT).
 Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 Feldmann K.;
 "Full-length cDNA from Arabidopsis thaliana";
 RT
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 [4]
 SEQUENCE FROM N.A. (ISOFORM LONG).
 STRAIN=cv. Columbia;
 Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 SSP consortium (Salk/Stanford/PGECC).";
 RT
 Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC
 -I- ALTERNATIVE PRODUCTS:
 CC
 Event=Alternative splicing; Named isoforms=2;
 CC
 Name=Long;
 CC
 IsoId=O64636-1; Sequence=Displayed;
 CC
 Name=Short;
 CC
 IsoId=O64636-2; Sequence=VSP_000619, VSP_000620;
 CC
 -I- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 -I- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
 in positions 483; 497 and 506.
 CC

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by, and for commercial/
 entities requires a license agreement. See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch.
 CC

 EMBL; D78600; BAA28540.1; ALT_FRAME.
 DR
 EMBL; AC003680; AAC06157.2; -
 DR
 EMBL; AY085090; AAM61644.1; -
 DR
 EMBL; AY078939; AAL84945.1; -
 DR
 PIR; T00869; T00869.
 DR
 InterPro; IPR001128; Cytochrome_P450.
 DR
 Pfam; PF00067; P450; 1.
 DR
 PRINTS; PR00385; P450.
 DR
 PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR
 Oxidoreductase; Monooxygenase; Heme; Transmembrane; Multigene family;
 KW
 Alternative splicing.

RT colour.";
 RN Nature 366:276-279(1993).
 RC [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=cv. Falcon Blue; TISSUE=Petal;
 RA Ohbayashi M., Shimada Y., Nakano R., Kiyokawa S., Kikuchi Y.;
 RT "Molecular cloning of cDNA encoding flavonoid-3',5'-hydroxylase from
 RT Petunia hybrida.";
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Hwang T.S., Chen Y.C., Wu R.Y.;
 RT "Cloning and sequencing of the genomic flavonoid 3',5'-hydroxylase
 RT (F3',5'Hase) gene from Petunia hybrida.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-425 FROM N.A.
 RA STRAIN=cv. Blue-star; TISSUE=flower;
 RA Toguri T., Azuma M., Ohtani T.;
 RT "The cloning and characterization of a cDNA encoding a cytochrome P450
 RT from the flowers of Petunia hybrida.";
 RL Plant Sci. 94:119-126(1993).
 CC -!- FUNCTION: Catalyzes the 3',5'-hydroxylation of naringenin and
 CC eriodictyol to form 5,7,3',4',5'-pentahydroxyflavanone and 3',5'-
 CC hydroxylation of dihydrokaempferol and dihydroquercetin to form
 CC dihydromyricetin.
 CC -!- PATHWAY: Anthocyanin biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Microsomal (Potential).
 CC -!- TISSUE SPECIFICITY: FLOWERS.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D22545; CAA80266.1; -;
 DR EMBL; D14588; BAA03438.1; -;
 DR EMBL; AF081575; AAC32274.1; -;
 DR EMBL; X71130; CAA50442.1; -;
 DR PIR; S32110; S32110.
 DR PIR; S38985; S38985.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Heme; Multigene family.
 FT METAL 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 506 AA; 56548 MW; FCF3F9623BD82B3E CRC64;
 Query Match 29.2%; Score 756; DB 1; Length 506;
 Best Local Similarity 32.1%; Pred. No. 1.8e-36;
 Matches 159; Conservative 114; Mismatches 200; Indels 22; Gaps 5;
 QY 10 LFLVLFLGVKVFQSRKLRNIPPGPPPLIIGNLNLLRQPIHRFFQRMKQYGNVSLWF 69
 DB 13 IFLIAHIIISTLSKTTGRHLPGPGRGWPGVIGALPLLGMFVSLAKWAKYGAIMYLVK 72
 QY 70 GSRLAVVSSPTAQECCTKHVALANRLPSLGKYPYNTTVGSCSHGHEWNLARIT 129
 DB 73 GTCGNMAVASTPDAKAFUKTLTDINFSRPPNAGATHLAYNAQDMVFAHYGPRWKLRLKS 132
 QY 130 ALDVLSTORVHSFGISRDETKRLMQRLVLAKNSEEFARVEISSMENDLTYNIMEMI 189
 DB 133 NLHMGKALENANVRANRELGHMKSM-----SDMSREGQVVVAEMLTFTFAMANNIGQVM 188
 QY 190 SKGRFYGESESMKNVEEAREFETVTEMLMGLANKGDHLPFLRWFDFFQNVKRLKXIS 249
 DB 189 LSKRFVVDKG-----VEVNEFKDMVVELMTIAGYFNIGDFIPCLAWDLQGIKRMKRLH 243

QY 250 KRYDSILKILHNENRASNDQRQNSMDHLKLOE-----TOPQYVTDQIIKGLALAMFGCT 305
 DB 244 KKFDFALLTRFDEHKATTVYERKGFDFLDVVMENGDSERGLSTTNIKALLINLFTAGT 303
 QY 306 DSTGTGLEWLSNLNHNPELVAKKARDELDTQVGDRLNLSLPLKLPYLRKIILETLRLY 365
 DB 304 DTSSSAIEWALAEWMKNPAILKKAQAEQDVIGRRRLLESIPNLPYLRACKETFRKH 363
 QY 366 PPAPILIPHVSSEDTIEGFNIPRDTIVINGMGHQRDPQLWNDATCFKPERP----- 418
 DB 364 PSTPLNLPRISEPCIVDGYIPIKNTLSVNIWALGRDPQVWENPLEFNPFRFLSGRNSK 423
 QY 419 -DVEGEKKLVAFGMRACPCGPEPMAMQSVSTGLGIIOCFQWKVSEK-KLDMTENWNI 476
 DB 424 IDPRGNDPELIIPFGAGRRICAGTRMGIVNVVETILGTLVHSEFDWKLPSEVIELNMEEAFGL 483
 QY 477 TLSRLIPLEAMCKAR 491
 DB 484 ALQKAVPLEAMVTPR 498
 RESULT 10
 CP71 PERAE
 ID CP71 PERAE STANDARD; PRT; 471 AA.
 AC P24465;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 71A1 (EC 1.14.-.-) (CYPLXXIAI) (ARP-2).
 GN CYP71A1.
 OS Persea americana (Avocado).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Persea.
 OX NCBI_TaxID=3435;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Hass;
 RX MEDLINE=90251665; PubMed=1692626;
 RA Bozak K.R., Yu H., Sirevag R., Christoffersen R.E.;
 RT "Sequence analysis of ripening-related cytochrome P-450 cDNAs from
 RT avocado fruit.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3904-3908(1990).
 RN [2]
 RP SEQUENCE OF 1-40.
 RC TISSUE=Mesocarp;
 RA O'Keefe D.P., Leto K.J.;
 RT "Cytochrome P-450 from the mesocarp of avocado (Persea americana).";
 RL Plant Physiol. 89:1141-1149(1989).
 CC -!- FUNCTION: INVOLVED IN THE METABOLISM OF COMPOUNDS ASSOCIATED WITH
 CC THE DEVELOPMENT OF FLAVOR IN THE RIPENING FRUIT PROCESS, POSSIBLY
 CC BY ACTING AS TRANS-CINNAMIC ACID 4-HYDROLASE.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: MICROSOMAL; MEMBRANE-BOUND (POTENTIAL).
 CC -!- TISSUE SPECIFICITY: MESOCARP.
 CC -!- DEVELOPMENTAL STAGE: IN RIPENING FRUIT.
 CC -!- INDUCTION: By ethylene.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M32885; AAA22913.1; ALT_TERM.
 DR PIR; A35867; A35867.
 DR HSPSP; P14779; IJJPZ.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.


```

C753 PETHY
ID -C753 PETHY STANDARD; PRT; 508 AA.
AC P48419;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE Flavonoid 3',5'-hydroxylase 2 (EC 1.14.-.-) (F3',5'H) (Cytochrome P450
DE 75A3) (CYPLXXV33).
GN CYP75A3 OR HF2.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanaceae; Solanaceae; Petunia.
OX NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Old Glory Blue; TISSUE=Petal;
RX MEDLINE=94050182; PubMed=8232589;
RA Holton T.A., Brugliera F., Lester D.R., Tanaka Y., Hyland C.D.,
RA Menting J.G.T., Lu C.-Y., Farcy E., Stevenson T.W., Cornish E.C.,
RT "Cloning and expression of cytochrome P450 genes controlling flower
RT colour."
RL Nature 366:276-279(1993).
CC -1- FUNCTION: Catalyzes the 3',5'-hydroxylation of naringenin and
CC eriodictyol to form 5,7,3',4',5'-pentahydroxyflavanone and 3',5'-
CC hydroxylation of dihydrokaempferol and dihydroquercetin to form
CC dihydromyricetin.
CC -1- PATHWAY: Anthocyanin biosynthesis.
CC -1- SUBCELLULAR LOCATION: Microsomal (Potential).
CC -1- TISSUE SPECIFICITY: FLOWERS.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 222544; CAA80265.1; -.
CC PIR; S38984; S38984.
CC InterPro; IPR001128; Cytochrome_P450.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Heme; Multigene family.
CC METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC FT METAL 508 AA; 56721 MW; 70989PF2B3AD4FD8 CRC64;
CC SEQUENCE
CC
CC Query Match 28.5%; Score 738.5; DB 1; Length 508;
CC Best Local Similarity 31.8%; Pred. No. 1.8e-35;
CC Matches 161; Conservative 110; Mismatches 206; Indels 29; Gaps 6;
CC
CC 9 VLFVLFLGVKVFQSRKLNIPGCPPLPIIGNLNLEQIHRFFORMSKQYGNVSLW 68
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 12 LIFLTHFIFLTSITNGRLPPGPGWPGVIGALPLGAMPVHVSIAKMAKYGAIMYLK 71
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 69 FGSRLAVVISPTAYQECFTKHDVALANRLSLGSKYIFNNYTTVSGSCHEHNRRLRI 128
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 72 VGTGCMVVASTPDAAKAFKLTLDLNFSPNPNAGATHLAYGAQDMVFHAPRKLKRL 131
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 129 TALDVLSTQRVSHSGIRSDTKRLMQRLVLAKNSNEEFARVETISSMFDNTNNIMRM 188
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 132 SNLHMLGGKALENANVANRELGHMLKSMFDMSEGE---RVVVAEMLTAMANNICQV 187
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 189 ISGKRFPYGESEMKNVAEAREFRETVTTEMLMGLANKGDHLPFLRPFDFQNVKRLKSI 248
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 188 ILSKRVF-----VNVKGVEVNEFKDMVVELMTTAGYFNIGDFIPCLAWMDLQIEKGMKRL 242
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 249 SKRYDSILKILHENRASNDQNSMIDHLKLQETQPOYYTDQI-----IKGLALAMLPFG 304
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :
CC 243 HKKFDALTKMFDEHKATSYERKKGKPDFLDCVMENRDNSGERLSTTNIKALLNLNLTAG 302
CC :||| : : : : : : : : : : : : : : : : : : : : : : : : :

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:28:34 ; Search time 102 Seconds
(without alignments)
1262.434 Million cell updates/sec

Title: US-09-868-546A-2

Perfect score: 2592
Sequence: 1 LLVWSYAVLVFLVFLGVKFP.....RLIPLEAMCKARPLATKIGI 499

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL23:

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_mhc:
- 8: sp_organelle:
- 9: sp_phage:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:
- 15: sp_rvirus:
- 16: sp_bacteriap:
- 17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1671.5	64.5	499	Q9ZWF2	Q9zwf2 glycyrrhiza
2	1646.5	63.5	498	Q9ZFW6	Q9zfw6 cicer ariet
3	1646.5	63.5	498	Q9XFX0	Q9xfx0 cicer ariet
4	1581	61.0	499	Q9MBE4	Q9mbe4 lotus japon
5	1350	52.1	499	Q9XFX1	Q9xfx1 cicer ariet
6	1287.5	49.7	500	Q23154	Q23154 arabidopsis
7	1285	49.6	499	Q9SZT6	Q9szt6 arabidopsis
8	1276	49.2	509	Q9LHA1	Q9lha1 arabidopsis
9	1241	47.9	497	Q9SZT7	Q9szt7 arabidopsis
10	1229	47.4	492	Q23155	Q23155 arabidopsis
11	1227	47.3	494	Q8LCN8	Q8lcn8 arabidopsis
12	1219.5	47.0	495	Q23156	Q23156 arabidopsis
13	1212	46.8	543	Q22188	Q22188 arabidopsis
14	1211	46.7	501	Q8LA85	Q8la85 arabidopsis
15	1205	46.5	515	Q22185	Q22185 arabidopsis
16	1199.5	46.3	481	Q9C714	Q9c714 arabidopsis

17	1189.5	45.9	491	10	Q9LVD6	Q9lvd6 arabidopsis
18	1188.5	45.9	500	10	Q8LBE7	Q8lbe7 arabidopsis
19	1171.5	45.2	500	10	Q8S9L8	Q8s9l8 arabidopsis
20	1168.5	45.1	518	10	Q9SW67	Q9sw67 arabidopsis
21	1144.5	44.2	484	10	Q8H137	Q8h137 arabidopsis
22	1132	43.7	520	10	Q6S815	Q6s815 helianthus
23	1125	43.4	501	10	Q9SZU1	Q9szu1 arabidopsis
24	1078	41.6	496	10	Q9FN20	Q9fn20 arabidopsis
25	1078	41.6	507	10	Q8GZ20	Q8gz20 arabidopsis
26	1023	39.5	457	10	Q8VVA6	Q8vva6 arabidopsis
27	1019	39.3	732	10	Q94HA5	Q94ha5 oryza sativ
28	1012	39.0	518	10	Q94HA3	Q94ha3 oryza sativ
29	1000	38.6	517	10	Q9ATV6	Q9atv6 lolium rigi
30	999.5	38.6	527	10	Q94HA4	Q94ha4 oryza sativ
31	995.5	38.4	512	10	Q94HA6	Q94ha6 oryza sativ
32	993	38.3	517	10	Q9ATV5	Q9atv5 lolium rigi
33	983	37.9	517	10	Q9ATV4	Q9atv4 lolium rigi
34	979	37.8	414	10	Q940P7	Q940p7 arabidopsis
35	903.5	34.9	561	10	Q9LXB2	Q9lxb2 arabidopsis
36	888.5	34.3	500	10	Q9LXB3	Q9lxb3 arabidopsis
37	840	32.4	372	10	Q81653	Q81653 hemerocalli
38	828.5	32.0	513	10	Q9FPN2	Q9fpn2 matthiola i
39	824	31.8	540	10	Q9FVK6	Q9fvk6 pisum sativ
40	810.5	31.3	509	10	Q94FM3	Q94fm3 nicotiana t
41	808.5	31.2	523	10	Q49394	Q49394 arabidopsis
42	801	30.9	524	10	Q9SZ46	Q9sz46 arabidopsis
43	793.5	30.6	526	10	Q8S845	Q8s845 oryza sativ
44	783	30.2	338	10	Q9SZU0	Q9szu0 arabidopsis
45	782.5	30.2	240	10	Q9SML2	Q9sml2 cicer ariet

ALIGNMENTS

RESULT 1

Q9ZWF2 ID Q9ZWF2 PRELIMINARY; PRT; 499 AA.

AC Q1-MAY-1999 (TREMREL. 10, Created)

DT 01-MAY-1999 (TREMREL. 10, Last sequence update)

DT 01-MAY-2003 (TREMREL. 23, Last annotation update)

DE Cytochrome P450.

GN CYP GE-31.

OS Glycyrrhiza echinata (Licorice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrrhiza.

OX NCBI_TaxID=46348;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=2008263; PubMed=10540749;

RA Nakamura K., Akashi T., Aoki T., Kawaguchi K., Ayabe S.;

RT "Induction of isoflavonoid and retrochalcone branches of the flavonoid

RT pathway in cultured glycyrrhiza echinata cells treated with yeast

RT extract.";

RL Biosci. Biotechnol. Biochem. 63:1618-1620(1999).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

DR EMBL; AB022732; BAA74465.1; -

DR InterPro; IPR001128; Cytochrome_P450.

DR Pfam; PF00067; P450; 1.

DR PRINTS; PR00385; P450.

DR PROSITE; PS00086; CYTOCHROME P450; 1.

KW Heme; Monooxygenase; Oxidoreductase.

SQ SEQUENCE 499 AA; 57227 MW; 0F65DA06F88D3D64 CRC64;

Query Match 64.5%; Score 1671.5; DB 10; Length 499;

Best Local Similarity 61.2%; Pred. No. 1.5e-109;

Matches 304; Conservative 99; Mismatches 91; Indels 3; Gaps 1;

QY 1 LLVWSYAVLVFLVFLGVKFPQSRKRNIPGPPPLPIIGNLNLEQPIHFFORMSKO 60

3 ILLSLSYSVYVLLAFIFINIVIRARKFKNLPGPSPSLPIIGNLHLKRLHRTFKLSEK 62

QY 61 YGNVSLWFGSLRLAVISSPAYQECFTKNDVLANRPLSLGKYIFYNNTTGVSCSHGE 120
 Db 63 YGHVSLWFGSLRLVWSSASEFOCFTKNDVLANRPLSLGKYIFYNNTTGLSTSYGE 122
 QY 121 HWRNLRRITADVSTORVHSGIRSDGTRKLMQRLVLANRNEEESFARVEISSMENDL 180
 Db 123 HWRNLRRITADVLSNHRKINSFGIRSDGTRKLMQRLVLANRNEEESFARVEISSMENDL 179
 QY 181 TYNMIRMTISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWFDQ 240
 Db 180 TFNNIMRMISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWFDQ 239
 QY 241 VEKRLKISIKRYDLSILNKILHENRASNDRQNSMIDHLLKLOETOPOYYTDOIIKGLALAM 300
 Db 240 LEKRLKIDISIKRYDLSILNKILHENRASNDRQNSMIDHLLKLOETOPOYYTDOIIKGLALAM 299
 QY 301 LFGGTDSSGTLEWSLSNLNHPVLKARDELDTQVQDRLNLSDELPLKLYLKIILE 360
 Db 300 LLAGTDSASVTLEWSMSNLNHPVLKARDELDTQVQDRLNLSDELPLKLYLKIILE 359
 QY 361 TRLYPAPILIPHVSSSEDITIEGNIPTDITVINGMQRDPOLWMDATCFKPERFDV 420
 Db 360 TRLYTAPALLPHSTSDCCNIGYKVPDITVILINAWAHRDPDLWTEATTFKPERFEK 419
 QY 421 EGEEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDQKRVSEKLDNTENNWTLSR 480
 Db 420 KGELEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDQKRVSEKLDNTENNWTLSR 479
 QY 481 LIPLEAMCKARPLATKI 497
 Db 480 LVPLKAMCKSRPVNKV 496

RESULT 2

Q9ZRW6 PRELIMINARY; PRT; 498 AA.
 AC Q9ZRW6;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cytochrome P450 (EC 1.14.14.1).
 GN CYP81E3.
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
 OX NCBI_TaxID=3827;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. ILC3279;
 RA Overkamp S., Barz W.;
 RT "Cloning of two Cicer arietinum L. cDNA's encoding Cytochrome P450s
 RT highly homologous to Isoflavone 2'-Hydroxylase from Licorice.";
 RL Plant Physiol. 120:935-935(1999).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AJ012581; CAA10067.1; --
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 498 AA; 57397 MW; 04BDC8E8EBF061D1 CRC64;

Query Match 63.5%; Score 1646.5; DB 10; Length 498;
 Best Local Similarity 59.4%; Pred. No. 8.5e-108;
 Matches 295; Conservative 101; Mismatches 98; Indels 3; Gaps 1;
 QY 1 LLVVSVAVLFLVFLGKVFQSKRLNIPPGPPPIIIGNLLEQPIHRFFORMSKQ 60
 Db 3 ILSLCSLFLVLSFFIIRLLFQSKRKNLPPGPPSPPIIGNLHLKRLHRTFKGLSKT 62
 QY 61 YGNVSLWFGSLRLAVISSPAYQECFTKNDVLANRPLSLGKYIFYNNTTGVSCSHGE 120

Db 63 YGDIIISLWFGSLRLVWVSSLSSEFOCFTKNDVLANRPRFLSGKYIFYNNTTGLSTSYSE 122
 QY 121 HWRNLRRITADVSTORVHSGIRSDGTRKLMQRLVLANRNEEESFARVEISSMENDL 180
 Db 123 HWRNLRRITADVLSNHRKINSFGIRSDGTRKLMQRLVLANRNEEESFARVEISSMENDL 179
 QY 181 TYNMIRMTISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWFDQ 240
 Db 180 TFNNIMRMISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWFDQ 239
 QY 241 VEKRLKISIKRYDLSILNKILHENRASNDRQNSMIDHLLKLOETOPOYYTDOIIKGLALAM 300
 Db 240 LEKRVANIDAKTDFAIRGLLQQRKCKERTNTMIDHLLMLQESQPEYYTDOIIKGLALAM 299
 QY 301 LFGGTDSSGTLEWSLSNLNHPVLKARDELDTQVQDRLNLSDELPLKLYLKIILE 360
 Db 300 LLAGTDSASVTLEWSMSNLNHPVLKARDELDTQVQDRLVDESDELPLKLYLKNVIE 359
 QY 361 TRLYPAPILIPHVSSSEDITIEGNIPTDITVINGMQRDPOLWMDATCFKPERFDV 420
 Db 360 TRLCTPAPALLPHSTSDCCNIGYKVPDITVILINAWAHRDPKSWSEATSFKPERFDK 419
 QY 421 EGEEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDQKRVSEKLDNTENNWTLSR 480
 Db 420 KGEIEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDQKRVSEKLDNTENNWTLSR 479
 QY 481 LIPLEAMCKARPLATKI 497
 Db 480 LIPLKAMCKTRPVNKV 496

RESULT 3

Q9XFX0 PRELIMINARY; PRT; 498 AA.
 AC Q9XFX0;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cytochrome P450 monooxygenase (EC 1.14.14.1).
 GN CYP81E3V2.
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
 OX NCBI_TaxID=3827;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. ILC 3279;
 RA Overkamp S., Barz W.;
 RT "Cloning of two Cicer arietinum L. cDNA's encoding Cytochrome P450s
 RT highly homologous to Isoflavone 2'-Hydroxylase from Licorice.";
 RL Plant Physiol. 120:935-935(1999).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AJ238439; CAB41490.1; --
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 498 AA; 57524 MW; F26675C98F771BD8 CRC64;

Query Match 63.5%; Score 1646.5; DB 10; Length 498;
 Best Local Similarity 60.4%; Pred. No. 8.5e-108;
 Matches 297; Conservative 96; Mismatches 96; Indels 3; Gaps 1;
 QY 6 SYAVLFLVFLGKVFQSKRLNIPPGPPPIIIGNLLEQPIHRFFORMSKQGNV 65
 Db 8 AYSLFLVLSFFIIRLLFQSKRKNLPPGPPSPPIIGNLHLKRLHRTFKGLSKTGDII 67
 QY 66 SLWFGSLRLAVISSPAYQECFTKNDVLANRPLSLGKYIFYNNTTGVSCSHGEHWRNL 125
 Db 68 SLWFGSLRLVWVSSLSSEFOCFTKNDVLANRPRFLSGKYIFYNNTTGLSTSYSEHWRNL 127

QY 126 RRITADVLTQVHVSFSGIRSDTKRLMORVLAKNSNEEFAVEISSMENDLTNNI 185
 DB 128 RRITSDVLSNHRINNFSGYRRDETQRLITKLA---EDSSTPAEVELSFRFYDMTFNNI 184
 QY 186 MRISGRFYGESESEKMNVEAREFRETTEMLMGLANKGHDLPFLRWFDFQNVKRLK 245
 DB 185 MRISGRFYGESESEKMNVEAREFRETTEMLMGLANKGHDLPFLRWFDFQNVKRLK 244
 QY 246 KSISKRYDSILNKILHENRASNDRONSMIDHLLKLOETQPOYYTDOIIKGLALAMLFGGT 305
 DB 245 KNIADKTDAFLRGLLQEQRNKERTNTMIDHLLNQESQPEYYTDOIIKGLALAMLLACT 304
 QY 306 DSSTGTLEWSLNLNHPVLKARDELDTQVQDRLLNESDLPKPYLRKIILTLRLY 365
 DB 305 DSSAVTLEWSMNLNHPVLKVRDELDTQVQDRLLNESDLPKPYLRKIILTLRLY 364
 QY 366 PPAPILIPHVSSEDTIEGNIPTDTIIVINGMORDPOLWMDATCFKPERFVGESEK 425
 DB 365 TPAPILIPHVSSEDDCTIGYKVPRTIVLINAWAHRDPKSWSEATTFKPERFVGESEK 424
 QY 426 KLVAFGMRGRACPGEPMAQSVFTLGLLIQCFDWKRVSEBKLDMTENNWTLSRLIPLE 485
 DB 425 KVIAFGMRGRVCPGEALALRTISMTLALLIQCDFWKRTSDMDMAERDGFVLKLVPLK 484
 QY 486 AMCKARPLATKI 497
 DB 485 AMCKTRPVVVKI 496

RESULT 4
 Q9MBE4 PRELIMINARY; PRT; 499 AA.
 ID Q9MBE4
 AC Q9MBE4
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Cytochrome P450.
 OS Lotus japonicus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
 OX NCBI_TaxID=34305;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Root;
 RC PubMed=11164575;
 RX Shimada N., Akashi T., Aoki T., Ayabe S.;
 RA "Induction of isoflavonoid pathway in the model legume Lotus
 RT japonicus: molecular characterization of enzymes involved in
 RL phytoalexin biosynthesis."
 RL Plant Sci. 160:37-47(2000).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AB025016; BAA93634.1; -.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 499 AA; 57357 MW; 59FD1DBD6DA3353D CRC64;

Query Match 61.0%; Score 1581; DB 10; Length 499;
 Best Local Similarity 59.8%; Pred. No. 3.5e-103;
 Matches 294; Conservative 90; Mismatches 104; Indels 4; Gaps 2;
 QY 7 YAVLFVLFLGVFVFSQSRKLNIPGPPPLPIIGNLNLEQPIHFRFFQMSKQYGNVVS 66
 DB 9 YSFYVALFAIKFLGSRKFNLPDPTSLPIIGNLHUKRPHLHFRFALSKYGDVFS 69
 QY 67 LWFSGRLAVVISPTAYQECFTKHDVALANRLPSLSGKIYFYNNNTTVGSCSHGEHWRNLR 126
 DB 69 LWFENRLVVVVSSFADVQECFTKNDVVLNANRFLSGKIYFYNTTLGTSYGEHWRNLR 128
 QY 127 RITALDVLSTQVHVSFSGIRSDTKRLMORVLAKNSNEEFAVEISSMENDLTNNI 186

DB 129 RITSLDVLNHRINFSPIRDETTLRLIRKLA---EDSAKNFSEVELTSRFFDMTFNNIM 185
 QY 187 RMISGRFYGESESEKMNVEAREFRETTEMLMGLANKGHDLPFLRWFDFQNVKRLK 246
 DB 186 RMISGRFYGESESEKMNVEAREFRETTEMLMGLANKGHDLPFLRWFDFQNVKRLK 245
 QY 247 SISIKRYDSILNKILHENRASNDRONSMIDHLLKLOETQPOYYTDOIIKGLALAMLFGGT 305
 DB 246 GISSKTDRLRGLLQEQRNKERTNTMIDHLLNQESQPEYYTDOIIKGLALAMLLACT 305
 QY 306 DSSTGTLEWSLNLNHPVLKARDELDTQVQDRLLNESDLPKPYLRKIILTLRLY 365
 DB 306 DSSAVTLEWSMNLNHPVLKVRDELDTQVQDRLLNESDLPKPYLRKIILTLRLY 365
 QY 366 PPAPILIPHVSSEDTIEGNIPTDTIIVINGMORDPOLWMDATCFKPERFVGESEK 425
 DB 365 TPAPILIPHVSSEDDCTIGYKVPRTIVLINAWAHRDPKSWSEATTFKPERFVGESEK 425
 QY 426 KLVAFGMRGRACPGEPMAQSVFTLGLLIQCFDWKRVSEBKLDMTENNWTLSRLIPLE 485
 DB 426 KLIPFGLGRACPGELLAIKRAISMTLALLIQCDFWKRVSEBKLDMTENNWTLSRLIPLE 485
 QY 486 AMCKARPLATKI 497
 DB 485 AMCKTRPVVVKI 496

RESULT 5
 Q9XFX1 PRELIMINARY; PRT; 499 AA.
 ID Q9XFX1
 AC Q9XFX1
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Cytochrome P450 (EC 1.14.14.1).
 GN CYP81E2.
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 OX NCBI_TaxID=3827;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. ILC 3279;
 RC PubMed=10773344;
 RX Overkamp S., Hein F., Barz W.;
 RA "Cloning and characterization of eight cytochrome P450 cDNAs from
 RT chickpea (Cicer arietinum L.) cell suspension cultures."
 RL Plant Sci. 155:101-108(2000).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AJ239051; CAB43505.1; -.
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00225; CRYSTALLIN BETAGAMMA; 1.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 499 AA; 56656 MW; E61B3A1D49E5BF55 CRC64;

Query Match 52.1%; Score 1350; DB 10; Length 499;
 Best Local Similarity 50.6%; Pred. No. 6.6e-87;
 Matches 249; Conservative 113; Mismatches 124; Indels 6; Gaps 5;
 QY 9 VLFVLFLGVFVFSQSRKLNIPGPPPLPIIGNLNLEQPIHFRFFQMSKQYGNVVS 67
 DB 9 IIFSLFVLFLGVFVFSQSRKLNIPGPPPLPIIGNLNLEQPIHFRFFQMSKQYGNVVS 68
 QY 68 WFGSRLAVVISPTAYQECFTKHDVALANRLPSLSGKIYFYNNNTTVGSCSHGEHWRNLR 127
 DB 69 KFGSRLAVVISPTAYQECFTKNDVVLNANRFLSGKIYFYNTTLGTSYGEHWRNLR 128

```

QY 128 ITALDVLSTORVHSFGIRSDETKRLMORLVLAKNSENEEFARVEISSMFNDITYNNIMR 187
Db 129 ISSIEILSTHRLNSFSEIRKDETMRLIGK--LAKSH-KDFTKVELRSLFGLTFTIMR 185
QY 188 MISGRPYGESEKONVEAREPRETVTEMLMGLANKGDHLPFLRWFDFQNVKRLKS 247
Db 186 MVGCKRYGESDGTNAEAKFRDMNNEQEFGLSGLNDFVPLFKLFDPSGGHKLKK 245
QY 248 ISKRYGILNKLHNENASNDRO-NSMIDHLLKLOETPOPYTDQIIKGLALAMLFGGTD 306
Db 246 VGEKMDALFOGLVDEHRRNKKKLNLTWIDHLLSLOESOPYYSQIIKGLIMALIIVAGTE 305
QY 307 SSTGTLEWSNLNHNPEVLKKADELDTQVGDRLNENSDLPKLPYLRLKILLETLLRYP 366
Db 306 TSSVTLEWAMSNLHNPELLEKAKIELDNHIGOEHLIEEAATKQLYQNIISLTLRHP 365
QY 367 PAPILIPHVSSEDITIEGFNI PRDTIIVINGWGMQRPOLWNDATCFKPERFD-VSEGEK 425
Db 366 FVTMLLPHLSHDCTIGGYDVRNTMLVNAWAIHRDPNLWADPMSFKPERFENGQDIDG 425
QY 426 KLVAFGMGRACPGCEPMAMOSVFTLLGLLIQCFFDKRVSEKLDMTENNWITLSRLIPLE 485
Db 426 GFTPFGMRACPGSGALATLGLTIGSLIQCFFEWKRAIGKEEVDMSGSGTVPKAIPL 485
QY 486 AMCKARPLATKI 497
Db 486 AQCKARPIINKI 497

RESULT 6
O23154 ID O23154 PRELIMINARY; PRT; 500 AA.
AC O23154:
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cytochrome P450-like protein.
CYAL0.20 OR AT4G37340.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP Rose M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.,
RA Chalwatzis N.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Bevan M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; Z99707; CAB16769.1; -.
DR EMBL; AL161591; CAB80400.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; Cytochrome P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 5798 MW; F33F5D15A78A413F CRC64;

Query Match 49.7%; Score 1287.5; DB 10; Length 500;
Best Local Similarity 49.7%; Pred. No. 1.7e-82;
Matches 247; Conservative 106; Mismatches 129; Indels 15; Gaps 9;

QY 9 VLELVFLGVKVFV---QSRKLRNIPGPP-PLPIGNLNLLEOPIHRRFFORMSKQYGN- 63
Db 6 LIFTFLFISLTLTIGRIKRRPNLPFPSPSWALPVIGHRLRLKPLHRRVFLSVSESLGDA 65

```

```

QY 64 -VVSWMFGSRSLAVVISSPAYQECFTKHDVALANRLPSLSGKYIFYNNTTVGSCSHGEHW 122
Db 66 PIISLRNLGNLRFVVSLSLAEECFTKDVVLNANRFLNSLASCHISYGCTTVVTASYGDHW 125
QY 123 RNLRIITALDVLSTORVHSFGIRSDETKRLMORLVLAKNSENEEFARVEISSMFNDITY 182
Db 126 RNLRIIGAVEIFSARHLNSFSIRDEIHRLI--ACLSRNS-LEFTKVMKSMFSNLTFF 182
QY 183 NNIMMISGRPYGESEKONVEAREPRETVTEMLMGLANKGDHLPFLRWFDFQNV 242
Db 183 NNIMMISGRPYGESEKONVEAREPRETVTEMLMGLANKGDHLPFLRWFDFQNV 242
QY 243 KELKSISKRYGILNKLHNENASND-RQNSMIDHLLKLOETPOPYTDQIIKGLALAML 301
Db 239 KRKKIASRLDEFLQGLVDERREGKEKQNTWMDHLLCLQETQPEYTTDIIKGLMSLI 298
QY 302 FCGTSSSTGTLEWSNLNHNPEVLKKADELDTQVGDRLNENSDLPKLPYLRLKILLET 361
Db 299 LAGTDTSAVTLEWTLNHNPEVLKKADELDTQVGDRLNENSDLPKLPYLRLKILLET 358
QY 362 LRLYPAPILIPHVSSEDITIEGFNI PRDTIIVINGWGMQRPOLWNDATCFKPERFDVE 421
Db 359 LRLYPAPILIPHVSSEDITIEGFNI PRDTIIVINGWGMQRPOLWNDATCFKPERFDVE 418
QY 422 GEEKLVAFGMGRACPGCEPMAMOSVFTLLGLLIQCFFDKRVSEKLDMTENNWITLSR 480
Db 419 GEAKLVAFGMGRACPGCEPMAMOSVFTLLGLLIQCFFDKRVSEKLDMTENNWITLSR 478
QY 481 LIPLAMCKARPLATKI 497
Db 479 AIPLVAMCKARPVVGKI 495

RESULT 7
Q9SZT6 ID Q9SZT6 PRELIMINARY; PRT; 499 AA.
AC Q9SZT6:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cytochrome P450-like protein.
FEGL17.10 OR AT4G37360.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP Rose M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.,
RA Bevan M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP Rose M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AL035601; CAB38203.1; -.
DR EMBL; AL161591; CAB80400.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; Cytochrome P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.

```


DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 495 AA; 55882 MW; 083BEF6772741AF5 CRC64;

Query Match 47.08; Score 1219.5; DB 10; Length 495;
Best Local Similarity 47.34; Pred. No. 1e-77;
Matches 232; Conservative 100; Mismatches 147; Indels 11; Gaps 6;

QY 11 FLVLFLGVKVFV-OSRKLRIIPGPP-PLPIIGNLNLEQPIHFFORMSKQYCN--VVS 66
DB 11 FLFTISIKULLTKSNKPNLPSPAYPLPVIGHLLKQPVHRTFHSISKSLGNAPIFH 70
QY 67 LWFOSRLAVISSPTAYOEFTKHDVALANRLPSLSKYIFYNNTTVGSSGSHGHWNL 126
DB 71 LRLGNRLVVISSHSIAEECTKNDVVLNRPDIIMAKHGYNFTNMAAASGHWNL 130
QY 127 RITADVLTSTORVHSFGSIRSDTKRLMORVLAKNSNEEFARVEISSMFDNLTNNIM 186
DB 131 RIAAVEIFSSHRISTFSIRKDEIRRLITHL---SRDSLHGFEVELEKSLTLNLA FNII 187
QY 187 RMISGRFYGBESEMKNVEAREFRTVTMELMGLANKGDHLPFLRWFDQNVKRLK 246
DB 188 MWAGKRYGYGTG-NDNEAKVRELI AEIMAGAGSGNLADYLP SINW--VTNFENQTK 243
QY 247 SISRYUSILNKILHENRASNDRNSMIDHLLKLOTPQOYTDQIITKGLALAMFGT 306
DB 244 ILGNRLDRVLQKLVDEKRAEKEGQTLIDHLLSPQETEPEYTDVVIKGI LALVAGT 303
QY 307 SSTGLWLSNLNHPVLKARDELDTQVGDRLNEDSLPKLYRKIILETLLRYP 366
DB 304 TSSVTLWAMNSLNHPLEIKARAEIDDKIGSRLVEESDVNLVHLQIVSETLLRYP 363
QY 367 PAPILIPHVSSDITIEGFNIPRTIVIIINGWQMDPQLMNDATCFKPERFDVEGEKK 426
DB 364 AVPLLLPHFSSDECKVAGYDMPRTLLLTNNWAMHRDPLGWEERFKPERFEKEGE 423
QY 427 LVAFGMGRACPGPMAMQSVFTGLLIQCFDVKRWSEKLDNTENNWTLSRLIPL 486
DB 424 LMPFGMRRACPGALGKRLVSLALGCLIQSFEWERYGAELVDWTEGEGITMPKATPL 483
QY 487 MCKARPLATK 496
DB 484 MCKARAIVGK 493

RESULT 13
O22188 PRELIMINARY; PRT; 543 AA.
ID O22188
AC O22188;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Putative cytochrome p450.
GN AT2G23190
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vankken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana";
RL Nature 402:761-768 (1999).

RNA SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RL Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AC002391; AAB87112.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 543 AA; 62777 MW; 06332D6C090FDB12 CRC64;

Query Match 46.8%; Score 1212; DB 10; Length 543;
Best Local Similarity 47.1%; Pred. No. 4e-77;
Matches 235; Conservative 102; Mismatches 146; Indels 16; Gaps 9;

QY 3 VVVSVAVLVLVLGVKVFQSRKLR-NIPGPP-PLPIIGNLNLEQPIHFFORMSKQ 60
DB 50 LILSLAFLF---FISLKLFGKRHSKFNLPSPAPRPLPFIGHLLKQLPHRTFLFSQS 106
QY 61 YGN--VVSWFSRLAVISSPTAYOEFTKHDVALANRLPSLSKYIFYNNTTVGSCSH 118
DB 107 LGDAPISRLGNHLTVVSSYSIAEECTKNDIVLANRPKFLGHIEYNFTTWSAPY 166
QY 119 GEHWRNLRITADVLTSTORVHSFGSIRSDTKRLMORVLAKNSNEEFARVEISSM 178
DB 167 GDHWRNLRITADVLTSTORVHSFGSIRSDTKRLMORVLAKNSNEEFARVEISSM 223
QY 179 DITYNNIMMISGKPYGBESEMKNVEAREFRTVTMELMGLANKGDHLPFLRWFD 237
DB 224 DLTIINNLMVAGKPYGEGTEQDEV--ARRVQLIDEIVYRAGVNAADYIPILSWITD 281
QY 238 FQNVKRLKASISKRYDSILNKILHENRASNDRNSMIDHLLKLOTPQOYTDQIITKGLA 297
DB 282 F---EKGVKELASRVDFLOSLVDERRVHKQKNTWMDHLLSLOETOPYDITVLKGI 338
QY 298 LAMLFQGTDSSTGLWLSNLNHPVLKARDELDTQVGDRLNEDSLPKLYRKI 357
DB 339 IVMILAGTETAGTLEWAMNLNHPLEIKARAEIDDKIGSRLVEESDVNLVHLQIV 398
QY 358 ILETLLRYPAPILIPHVSSDITIEGFNIPRTIVIIINGWQMDPQLMNDATCFKPER 417
DB 399 VLETLLRYPVAPTNPHTSDDCILAGYDVPGRSMLLVNVWMSHRDPSIWEAPMPKPER 458
QY 418 EDVGEKEKLVAFGMGRACPGPMAMQSVFTGLLIQCFDVKRWSEKLDNTENNWT 477
DB 459 FKNEKLNQKLSFGFGRRACPGVGLAHLRLMSIALGSMVQCFEWRIGEEYVDTREBP 518
QY 478 LSRILPLEAMCKARPLATK 496
DB 519 MRPATPLAMCKARPIVHK 537

RESULT 14
O22188 PRELIMINARY; PRT; 501 AA.
ID O22188
AC O22188;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Cytochrome p450 monooxygenase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Haas B.J., Volfovsky N., Trukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:21:13 ; Search time 87 Seconds

(without alignments)
910.397 Million cell updates/sec

Title: US-09-868-546A-2

Perfect score: 2592

Sequence: 1 LUVVSYAVFLVFLGVKFP.....RLIPLEAMCKARPLATKIGI 499

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_19Jun03.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2592	100.0	499	21	AAV96593 Soybean isoflavone
2	2214	85.4	494	21	AAV96594 Soybean isoflavone
3	1251.5	48.3	498	21	AAV950170 Arabidopsis thalia
4	1241	47.9	497	21	AAV950170 Arabidopsis thalia
5	1241	47.9	497	21	AAV950195 Arabidopsis thalia
6	1241	47.9	497	21	ABP81276 Arabidopsis thalia
7	1229	47.4	492	21	AAV950128 Arabidopsis thalia
8	1229	47.4	514	21	AAV950127 Arabidopsis thalia
9	1227	47.3	494	21	AAV96985 Arabidopsis thalia

10	1216	46.9	501	21	AAV950190 Arabidopsis thalia
11	1216	46.9	516	21	AAV950189 Arabidopsis thalia
12	1211	46.7	501	21	AAV950189 Arabidopsis thalia
13	1211	46.7	516	21	AAV950189 Arabidopsis thalia
14	1196.5	46.2	502	21	AAV950189 Arabidopsis thalia
15	1196.5	46.2	505	21	AAV950189 Arabidopsis thalia
16	1188.5	45.9	500	21	AAV950189 Arabidopsis thalia
17	1188.5	45.9	500	21	AAV950189 Arabidopsis thalia
18	1188.5	45.9	519	21	AAV950189 Arabidopsis thalia
19	1188.5	45.9	519	21	AAV950189 Arabidopsis thalia
20	1125	43.4	501	21	AAV950189 Arabidopsis thalia
21	1125	43.4	501	21	AAV950189 Arabidopsis thalia
22	1125	43.4	509	21	AAV950189 Arabidopsis thalia
23	1125	43.4	509	21	AAV950189 Arabidopsis thalia
24	1123	43.3	505	20	AAV950189 Arabidopsis thalia
25	1100.5	42.5	457	21	AAV950189 Arabidopsis thalia
26	1100.5	42.5	457	21	AAV950189 Arabidopsis thalia
27	1099	42.4	499	19	AAV950189 Arabidopsis thalia
28	1080	41.7	429	21	AAV950189 Arabidopsis thalia
29	1050.5	40.5	383	21	AAV950189 Arabidopsis thalia
30	1040	40.1	382	21	AAV950189 Arabidopsis thalia
31	1040	40.1	382	21	AAV950189 Arabidopsis thalia
32	1033	39.9	379	21	AAV950189 Arabidopsis thalia
33	921.5	35.6	514	22	AAV950189 Arabidopsis thalia
34	919	35.5	326	21	AAV950189 Arabidopsis thalia
35	908.5	35.1	325	21	AAV950189 Arabidopsis thalia
36	908.5	35.1	325	21	AAV950189 Arabidopsis thalia
37	901.5	34.8	322	21	AAV950189 Arabidopsis thalia
38	890.5	34.4	319	21	AAV950189 Arabidopsis thalia
39	861	33.2	347	21	AAV950189 Arabidopsis thalia
40	858	33.1	347	21	AAV950189 Arabidopsis thalia
41	850.5	32.8	513	18	AAV950189 Arabidopsis thalia
42	810.5	31.3	509	24	AAV950189 Arabidopsis thalia
43	808	31.2	347	21	AAV950189 Arabidopsis thalia
44	808	31.2	347	21	AAV950189 Arabidopsis thalia
45	790.5	30.5	508	18	AAV950189 Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAV96593

ID AAV96593 standard; Protein; 499 AA.

XX AC AAV96593;

XX DT 26-SEP-2000 (first entry)

XX XX Soybean isoflavone-2-hydroxylase.

DE KW Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;

KW flower colour; pollen tube; feeding deterrent; UV irradiation.

XX OS Glycine max.

XX PN WO200037652-A2.

XX PD 29-JUN-2000.

XX PF 20-DEC-1999; 99WO-US30337.

XX PR 21-DEC-1998; 98US-0113190.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;

XX DR WPI; 2000-442678/38.

XX DR N-PSDB; AAA29326.

XX PT New polynucleotide encoding flavonoid biosynthetic enzymes, useful for

PT producing transgenic plants and immunological screening of cDNA

PT libraries

Claim 10; Page 31-32; 36pp; English.

This enzyme, a plant (soybean) isoflavone-2-hydroxylase, was encoded by cDNA, isolated from clone sisc.pk005.n3. It was determined using the sequence of an isoflavone-2-hydroxylase encoded by a contig composed of clones sg1c.pk001.g17, sg2c.pk004.h7 and slf1.pk0034.g1. The cDNA sequences can be used for the recombinant production of the enzyme, to isolate homologues, to create transgenic plants and to provide probes for genetically and physically mapping genes and as markers for traits linked to the genes. The proteins can be used for immunological screening, in particular to raise antibodies against the enzymes. The enzyme and its gene are useful to study flavonol biosynthesis in plants and provide means to enhance or otherwise alter flavonol and anthocyanin biosynthesis. Flavonoids have diverse functions, such as co-pigments in flower colour, stimulation of pollen tube growth, pollinator attraction, and feeding deterrents and protection against UV irradiation in fruits and seeds.

Sequence 499 AA;

Query Match 100.0%; Score 2592; DB 21; Length 499;

Best Local Similarity 100.0%; Pred. No. 4.8e-226;

Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLVVSVAVLFLVLFLGVKVFQSRKLRNIPGPPPLPIIGNLLEQPIHRFFORMSKQ 60
 Db 1 LLVVSVAVLFLVLFLGVKVFQSRKLRNIPGPPPLPIIGNLLEQPIHRFFORMSKQ 60
 Qy 61 YGNVSWLWFGSLAVVSSPTAYQECFTKHDVALANRLPSLGKGYFYNNNTVSGSHGE 120
 Db 61 YGNVSWLWFGSLAVVSSPTAYQECFTKHDVALANRLPSLGKGYFYNNNTVSGSHGE 120
 Qy 121 HWRNLRITADVLTSTQRVHSGIRSDTKRLMORLVAKNSNEEFARVEISSMFNDL 180
 Db 121 HWRNLRITADVLTSTQRVHSGIRSDTKRLMORLVAKNSNEEFARVEISSMFNDL 180
 Qy 181 TYNMIMISGRFYGESEEMKNVEAREPRETVMLEMLGKNGDHLPLFLRWDFQFN 240
 Db 181 TYNMIMISGRFYGESEEMKNVEAREPRETVMLEMLGKNGDHLPLFLRWDFQFN 240
 Qy 241 VEKRLKSKRYDYLILNKLHNENRASNDRQNSMIDHLLKLQETQPYVYTDQIIKGLALAM 300
 Db 241 VEKRLKSKRYDYLILNKLHNENRASNDRQNSMIDHLLKLQETQPYVYTDQIIKGLALAM 300
 Qy 301 LFGCTDSSTGTLEWLSNLNHPVLKKADELDTQVGDRLINESDLPKLPYLKRIILE 360
 Db 301 LFGCTDSSTGTLEWLSNLNHPVLKKADELDTQVGDRLINESDLPKLPYLKRIILE 360
 Qy 361 TLRYPAPALIPHVSSEDIITIEGFNIPRDTIVINGMGORDPOLWMDATCFKPERFDV 420
 Db 361 TLRYPAPALIPHVSSEDIITIEGFNIPRDTIVINGMGORDPOLWMDATCFKPERFDV 420
 Qy 421 EGEKKLVAFNGRRACGPEPMAMQSVFTGLLIQCPDWKRVSEKLDMTENNWITLSR 480
 Db 421 EGEKKLVAFNGRRACGPEPMAMQSVFTGLLIQCPDWKRVSEKLDMTENNWITLSR 480
 Qy 481 LIPLEAMCKARPLATKIGI 499
 Db 481 LIPLEAMCKARPLATKIGI 499

RESULT 2

AA96594
 ID AA96594 standard; Protein; 494 AA.

XX
 AC AA96594;

XX
 DT 26-SEP-2000 (first entry)

XX
 DE Soybean isoflavone-2-hydroxylase.

XX

KW Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;
 flower colour; pollen tube; feeding deterrent; UV irradiation.
 XX Glycine max.
 OS
 XX WO200037652-A2.
 PN
 XX 29-JUN-2000.
 PD
 XX 20-DEC-1999; 99WO-US30337.
 PF
 XX 21-DEC-1998; 98US-0113190.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;
 PI
 XX WPI; 2000-442678/38.
 DR
 XX N-PSDB; AAA29327.
 XX
 PT New polynucleotide encoding flavonoid biosynthetic enzymes, useful for
 producing transgenic plants and immunological screening of cDNA
 libraries
 PS
 XX Claim 10; Page 33-34; 36pp; English.
 XX
 CC This enzyme, a plant (soybean) isoflavone-2-hydroxylase, was encoded by
 cDNA, isolated from clone src3c.pk005.f5. It was determined using the
 sequence of an isoflavone-2-hydroxylase encoded by a contig composed of
 clones sg1c.pk001.g17, sg2c.pk004.h7 and slf1.pk0034.g1. The cDNA
 sequences can be used for the recombinant production of the enzyme, to
 isolate homologues, to create transgenic plants and to provide probes for
 genetically and physically mapping genes and as markers for traits linked
 to the genes. The proteins can be used for immunological screening, in
 particular to raise antibodies against the enzymes. The enzyme and its
 gene are useful to study flavonol biosynthesis in plants and provide
 means to enhance or otherwise alter flavonol and anthocyanin
 biosynthesis. Flavonoids have diverse functions, such as co-pigments in
 flower colour, stimulation of pollen tube growth, pollinator attraction,
 and feeding deterrents and protection against UV irradiation in fruits
 and seeds.
 XX
 SQ Sequence 494 AA;
 Query Match 85.4%; Score 2214; DB 21; Length 494;
 Best Local Similarity 84.4%; Pred. No. 9.2e-192;
 Matches 417; Conservative 45; Mismatches 28; Indels 4; Gaps 3;
 Qy 5 VSYAVLFLVLFLGVKVFQ-SRKLRNIPGPPPLPIIGNLLEQPIHRFFORMSKOYGN 63
 Db 1 LSYLLSLVFFFTLKYLFQSRKVRNLPFGPTPLPIIGNLNLVEQPIHRFFORMSKOYGN 60
 Qy 64 VVSLWFGSLAVVSSPTAYQECFTKHDVALANRLPSLGKGYFYNNNTVSGSHGEHWR 123
 Db 61 IISLWFGSLRVVSSPTAYQECFTKHDVTLANRVSLSGKGYFYDNTTVSGSHGEHWR 120
 Qy 124 NLRITADVLTSTQRVHSGIRSDTKRLMORLVAKNSNEEFARVEISSMFNDLTYN 183
 Db 121 NLRITADVLTSTQRVHSGIRSDTKRLIHR--LARDSG-KDFARVEVTSKFDLTYN 177
 Qy 184 NIMRMISGRFYGESEEMKNVEAREPRETVMLEMLGKNGDHLPLFLRWDFQVNEK 243
 Db 178 NIMRMISGRFYGESEELANVEAREFRDVTNEMQLMGLANKGDHLPLFLRWDFQVNEK 237
 Qy 244 RKLSIKRYDYSILNKLHNENRASNDRQNSMIDHLLKLQETQPYVYTDQIIKGLALAMLF 303
 Db 238 RLKNSIKRYDYLILNKLHNENRASNDRQNSMIGHLLKLQETQPYVYTDQIIKGLALAMLF 297
 Qy 304 GTDSSGTGLEWLSNLNHPVLKKADELDTQVGDRLINESDLPKLPYLKRIILETLR 363
 Db 298 GTDSSGTGLEWLSNLVNDPEVLQKARDELDAQVGDRLINESDLPKLPYLKRIILETLR 357
 Qy 364 LYPPAPILIPHVSSEDIITIEGFNIPRDTIVINGMGORDPOLWMDATCFKPERFDVEGE 423


```

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.3%; Score 1251.5; DB 21; Length 498;
Best Local Similarity 48.3%; Pred. No. 1.9e-104;
Matches ,239; Conservative 110; Mismatches 133; Indels 13; Gaps 6;

QY 9 VLFVLFVGVKVFQSRKLR--NIPGPP-PLPIGNLNLLEQPIHRRPFQMSKQYGN- 63
DB 6 LIFSILFVLSLIYLIGLKKRKNLPPSPAWSLPVIGHRLRLKPKPIHRTFLSLSLNNA 65
QY 64 -VVSILWFGSRLAVISSPTAYQECFTKHDVALANRLPSLSCKYIFYNNTTVGSSCHGRHW 122
DB 66 PIFSRLGNRLVFNSSHSIAECEFKNDDVVLARNPNFILAKHVAYDTTIIAASYGDHW 125
QY 123 RNLRRITALDVLSTQVHSGIRSDETKRLMORLVLAKNSEEEFARVEISSMFDNTLY 182
DB 126 RNLRRIGSVFISNHLNLSFLSIRKDEIRLVFRL---SRNFSQEFVKVDMKMSLSLDTFF 182
QY 183 NNIMRMISGRFYGESEMKVVEAREPRETVTEMLEMLGLANKGDHLPFLRWFDFQVVE 242
DB 183 NNILEMWAGKRYYGDE--DDPEAKVRQIADVVACAGAGNAVLYLVLRL--VSDYE 238
QY 243 KRLKSIKRYDSILNKILHENRASNDRONSMDHLLKLOETQPOVYTTQIIKGLALAMLF 302
DB 239 TRVKIAGRLDEFLOGLVDEKREAKEKNTWIDHLLTLOESQPDYFTDRIIKGNMLSLIA 298
QY 303 GGTDSTGTLEWSLSNLLNHPVLKKADELDTQVGDRLNEDSLPKLPYLRKIILETL 362
DB 299 GGTDTTAVTLEWALSLLNPNPEVLNKADEIDRMIGVDRLLSESDIPNLPLYQNIVSETL 358
QY 363 RLYPPAPILIPHVSSEDTITIEGFNIPRDTIIVINGMGNORDPOLWNDATCEKPERFDEG 422
DB 359 RLYPAAPMLLPHVASKCKCKGVDMRGTMLLTNAWAIHRDPLLLWDDPTSFKEPERFEKRG 418
QY 423 EEKLVAFGMGRACPGEPMAQSVFTLGLLIQCFDWMKRVSEKLDNTENNWITLSRLI 482
DB 419 EAKLMPFGLGRRACPGSLQRLVTLGLSGLIQCFEWERIGEEVDNTEGGLTWPQAR 478
QY 483 PLEAMCKARPLATKI 497
DB 479 PLEAMCRARDVFGKI 493

RESULT 4
AAG32080
ID AAG32080 standard; Protein; 497 AA.
XX
AC AAG32080;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38634.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

```

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

```
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 47.98; Score 1241; DB 21; Length 497;
Best Local Similarity 47.98; Pred. No. 1.7e-103;
Matches 237; Conservative 111; Mismatches 133; Indels 14; Gaps 7;

QY 9 VLEVLVFLGVKVFQGRKLR--NIPGPP-PLPIGNNLMLLOPIHRRFFORMSKOYGN- 63
DB 6 LIFSILFVLVLSLYLICKLKKENLPSPSPAWSLPVIGHLLEKPPHRTFLSLSQSNKA 65
QY 64 -VVSLMFGSLAVVISPTAQCFCKHDVALANRPLPSLGKVIYFNNITVGSCHGEHW 122
DB 66 PIFSRLGNRLVFNSSHSIAEBCFTKNDVVLANRPNFILAKHAVDYTTMIAASYGDHW 125
QY 123 RNLRRITADVLTQRVHSGSIRSDTKLMORVLAKNSNEEERARVEISSMFDLTY 182
DB 126 RNLRRIGSVEIFGNHRLNSFLSRKDEIRLVFRL--SRNFSQEFVKVDMKMSLDLTF 182
QY 183 NNIMRMISGRKFGESEEMKNVEAREFRETVTMELEMLGLANKGDLPLPLRWFDQNV 242
DB 183 NNILEWAGKRYGCVGE--DDPEAKVRQLIADVVACAGNAVLYLPVRL--VSDYE 238
QY 243 KRLKSIKRVDSILNKILHNENRASNDRQNSMIDHLLKLOETQOYYTDQIIKGLAMLP 302
DB 239 TRVKLAGRLDEFLQGVDEKREAKEKGTMDIHLTLQESQPDYFTDRIKGNMLALIL 298
QY 303 GGTDSSTGLTLEMSNLNHPVLKKARDELTVQGDRLNLSDLPLKPLRKILLET 362
DB 299 AGTDTSAVTLEWALSNNHPDLVNLKARDEIDRKIGLDRLMDESISNLPYLNIVSET 358
QY 363 RLYPPAPILIPHYSSDITIEGFI PRDTTIVINGMGMORDPOLMNDATCFKPERFDVEG 422
DB 359 RLYPAAPMLLPHYASEDKVAGYDMPRTGILLTNVWAIHRDPQLMDDPMSFKPERKEG 418
QY 423 EEKLVAFGMRACPCGEPAMOSVSTGLLIQCDFWKVSEKLDMTENNITSRLLI 482
DB 419 EAQKLMFPGLGRACPCGSLAHLINLTGLSLIQCLEWEKIGEE-VDMSEKGVTMPKAK 477
QY 483 PLEAMCKARPLATKI 497
DB 478 PLEAMCRARPSVKI 492

RESULT 5
AAG50195
ID AAG50195 standard; Protein; 497 AA.
XX
AC AAG50195;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63582.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
```

```
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
```


PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-01334256.
PR 14-MAY-1999; 99US-01334218.
PR 14-MAY-1999; 99US-01334219.
PR 14-MAY-1999; 99US-01334221.
PR 14-MAY-1999; 99US-01334370.
PR 18-MAY-1999; 99US-01334768.
PR 19-MAY-1999; 99US-01334941.
PR 20-MAY-1999; 99US-01335124.
PR 21-MAY-1999; 99US-01335353.
PR 24-MAY-1999; 99US-01335629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139419.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 47.4%; Score 1229; DB 21; Length 492;			
Best Local Similarity 47.4%; Pred. No. 2.le-102;			
Matches 237; Conservative 107; Mismatches 138; Indels 18; Gaps 7;			
Qy	2	LVVSVYAVLF-LVLFGLVKFVFQSKRLRNIPPGPP-PLPIIGNLNLLLEQPIHRPFQRMKS	59
Db	6	ILILSFLSLFLAIFL----LTRSKRLNPPSPAISLPVIGHLLHLLKPLHRTFLSLK	61
Qy	60	QYGN--VVSLWFGRLAVISSPAYOCFTKHVALANRLPSLGSKYIFYNNTTVGSCS	117
Db	62	SIGNAPVFLRLGNRLVVISSRSIAEBCFTKNDVLANRPKFTISKHLGYNATYLLSAS	121
Qy	118	HGEHRLRRITALDVLSTQRFVSHSGIRSDTKELMQLRVLANKNSNEEFARVEISSMF	177
Db	122	YGDHRLRLRAAEIFSTHRLNSFLYRKDEIRRLSHL---SRDSLHGFVEVEMKTL	178
Qy	178	NLTNNIMRISGRFYGESEEMKNVVEAREFRETVTEMLEMLGLANKGDHLPFLRFWD	237
Db	179	TNLSNTTIRMAGKRYEGDN-----DDAKLVKNLVSEAVTSAGAGNPIDYLSILRW--	231
Qy	238	FQNVKRLKSISKYDSILNKILHENRASNDRONSMIDHLLKLOETQOYYTDOIIKGLA	297
Db	232	VSSYEKRIKILGNRFDTPQLKLVDEKKAKEGETMIDHLLALQDIQPDYVTVIIRGII	291
Qy	298	LAMLFGGTSDSTGLWLSNLNHPVLKARDELTOVCQDRLLNESDLPLPYLRKI	357
Db	292	LTLLIAGTDSVTLWAMSLLNHPPEILKARWEIDKVLGLDRLVDESIVNLSYLSQSI	351
Qy	358	ILETLRYPAPILIPHVSSEDITIEGNIPTDITVIINGMQRDPQLMNDATCFKPER	417
Db	352	VLETLRMPYAPVLLPLHLSSEDCKVGVDIPSGTMTNLTNAWMHRDPEVWEDPEIFKPER	411
Qy	418	FVVEGEEKLVAFGMRACPGEPNMQSVSFTGLLIQCDFMKRVSEELDMTENNWT	477
Db	412	FEKEGEAEKLISFGMRACPGAGLAHRLINQALGSLVQCPEWERVGEDFVDMTEKCAT	471
Qy	478	LSRLIPLEAMCKARPLATKI	497
Db	472	LPKAIPLRAMCKARSIVDKL	491
RESULT 8			
ID	AAG30127		
ID	AAG30127 standard; Protein; 514 AA.		
XX	AAG30127;		
XX	17-OCT-2000 (first entry)		
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 35962.		
DE	Protein identification; signal transduction pathway; metabolic pathway;		
XX	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
KW	Arabidopsis thaliana.		
XX			

XX	EP1033405-A2.	
XX	06-SEP-2000.	
XX	25-FEB-2000;	2000EP-0301439.
XX	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	28-APR-1999;	99US-0130891.
PR	30-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	04-MAY-1999;	99US-0132407.
PR	05-MAY-1999;	99US-0132484.
PR	06-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	11-MAY-1999;	99US-0132863.
PR	14-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

```
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.9%; Score 1216; DB 21; Length 501;
Best Local Similarity 47.1%; Pred. No. 3.2e-101;
Matches 235; Conservative 117; Mismatches 127; Indels 20; Gaps 11;

QY 9 VLFLVLF-LGVKVFQSKLR-NIPGPP-PLPIIGNLNLLEQPIHREFFORMSKOYGVV 65
Db 7 VLPLALFLAYLFFTSKTRNLPPSPYSLPIGHNLNLLAPPVHRLFHRUSKTHGP 66

QY 66 SLWFGSLAVVSISSAYOECT-KHDVALANRLPSLCKYFYNNYTTGSCSHGHEWRN 124
Db 67 SLQFGSRRVAVSISSSLAQCTQGNDIILSNRPFCLTAKYVAYNVTGTAPYGDHWRN 126

QY 125 LRRITADVLTQVHVSFGISDETKRLMQRLVLAKNSNEEFARVEISSMFDLTNN 184
Db 127 LRRICSLILSSNRLTNFLHRRKDETHRLTR--LSRDVYKKE---IELEPLLSDLTFN 180

QY 185 IMRLISGKGFYGEESKMKVGEAREPRETVTEMLGLMGLANKGDHLPFLRWDFONVEKR 244
Db 181 IVRMTGKRYGD--EVHNEEANVPKLVADINDCSGARHPGDYLPFMKMGF-GSFEKK 237

QY 245 LKSIKRYDSILNKLILHENRASNDRONSIMIDHLLKLOETQPOYTTDQITKGLALAMLFGG 304
Db 238 VKALAEAMDELLQRLLEECKROK-GNTVWNHLLSLQONEPEYTTDVTIKGLMGLMWIAG 296

QY 305 TDSSTGTLWSLNLNHLNDEVUKARDELDTQVQDRLLNESDLKPLPYLRKILLETURL 364
Db 297 TDTSAVTLWAMSSLNHPEALEKAKLEIDEKIGQERLIDEPDIANLPYLQNVSETFL 356

QY 365 YPPAPILIPHVSSEDTIEGFNIPRTTIVINGWGNQDRPOLWNDATCFKPERF----- 419
Db 357 YPAAPLPLVRSPTEDIKVGIDVPRGTVMVWNAWAHROPPELWNEPEKPFERNNGEGG 416

QY 420 VEGEE-KKLVAFGMRRACPGEPMAQSVSFTLGLLIQCFDMKRVSEEEKLDMTENNWIPL 478
Db 417 GRGEDVHKLMPFGNGRSCFGAGLGQKIYTLALGSLIQCFDMQKYNGEALDITETPGMAM 476

QY 479 SRLIPLEAMKARPLATKI 497
Db 477 RKXPLSALCQSRPMSKL 495

RESULT 11
AAG50189
ID AAG50189 standard; Protein; 516 AA.
XX
AC AAG50189;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63574.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
```

```
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
```

PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0147260.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149529.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152263.
 PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 46.9%; Score 1216; DB 21; Length 516;

Best Local Similarity 47.1%; Pred. No. 3.4e-101;

Matches 235; Conservative 117; Mismatches 127; Indels 20; Gaps 11;

QY 9 VLFVLF-LGVKVFQSRKLR-NIPPGPP-PILPIGNLILLOPPIHRRFFORMSKOYGVV 65
 DB 22 VLPLALFLAYKLFPTSKTRFNLPSPYSLPILGHHNLLKPPVHRLFHLRSKTHGP 81
 QY 66 SLWFGSLAVISSPTAYQECFT-KHDVALANRLPSLSGKYIFYNNITVVGSCSHGEHWRN 124
 DB 82 SLOFGRRAVVISSSLATQCTGNDIILNRPCLTAKYVAYNYTTGTAPYGDHWRN 141
 QY 125 LRRITADVLTQVHSPSGIRSDTKLMORLVLAKNSNEEPARVEISSMFDLTNN 184
 DB 142 LRRICSLILSSRLTNFLHFKDEIHRMLTR--LSRDVNKE---IELEPLLSLDTFNN 195
 QY 185 IMRMISGRFYGESEMKNVEEAREFRTVTEMLEMLGLANKGDHLPFLRWDFQVNEKR 244
 DB 196 IVRMVTGKRYGD--EVHNEEAVNFKGLVADINDCSGARHPGDYLPFMKMGF-GSFEKK 252
 QY 245 LKTSIKRYDSILNKILHENRASNDRONSMIDHLLKQETQPOYYTDQIIGLALAMFG 304
 DB 253 VKALAEAMDEILQRLLECKDKD-GNTVNHLLSLQNEPEYYTDTVIKGLMGMWAG 311
 QY 305 TDSSTGTTLEWSLNLNHPVLKKADELDTQVGDRLNLSDELKPLPYLRKILLETURL 364
 DB 312 TDTSVATLEWAMSSILNHPLEAKLBIDEKIGERLIDSPDIANLPYLNQIVSETRL 371
 QY 365 YPPAPILPHVSSSEDITIEGFNIPRDTIIVINGWGMQDDPOLWNDATCFKPERPD----- 419

```
Db 372 YPAAPLLVPRSTEDIKVGIDVPRGTMVWNAWAHHRDPELWNEPKFKPERFNGGEGG 431
Qy 420 VEGEE-KKLVAFQMGRRACPGSPMAMQSVSFTLGLLIQCFFDKWKRVSSEKLDMTENNWITL 478
Db 432 GRGEDVHKLMPGNGRRSCPGAGLGQKIVTLALGSLIQCFFDKWQKNGEALDWTETPGMAM 491
Qy 479 SRLIPLEAMCKARPLATKI 497
Db 492 RKKIPLSALCQSRPIMSIL 510

RESULT 12
AAG44324
ID AAG44324 standard; Protein; 501 AA.
XX AC AAG44324;
XX XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55508.
DE
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139432.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
```



```
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.2%; Score 1196.5; DB 21; Length 502;
Best Local Similarity 47.1%; Pred. No. 1.9e-99;
Matches 236; Conservative 100; Mismatches 150; Indels 15; Gaps 7;

QY 4 VVSVAFLVFLVGLGVKFVQSRKURNIPGPPPP-LPIIGNMLLEOPIHRPFQRMKQY- 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 8 VVLYSI-FSLIFLIISFKFLPKKQNLPPSPGWLPIIGHLLRLLKPPIHRTLSRFSETLD 66
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 62 ---GNVSLWFGSELAVISS-PTAYOCCTKHVALANLPSLSGKYIFVNNVTGSC 116
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 67 HNDGGVMSLRIGSLRVVSSHVAEECFGNKDVLANRPQVIGKHVGNNNTMIAA 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 117 SHGEHWRNLRLTALDVLSTORVHSFGIRSDTRKMLQRLVLAKNSNEEFARVEISSM 176
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 127 PYGDHWRNLRLCTIEIFSTRHLCNCLVVRTDEVRLISRLSRLAGTKK--TVVELKPM 183
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 177 FNDLYNNIMRMISCKRPFYGESEKMNVEAREFRETVTTEMLMGLANKGDHLPFLRWF 236
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 184 LMDLTFNNIMRMWTGKRYGEEET--TDEEAKRVRKLVADVAGANTSSGNAVDPVPLRL- 240
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 237 DFQNVKELKSIKRYDSILNKLHNENRASNDRQNSMDHLLKIQETOPQVYTDQIKGL 296
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 241 -FSSYENRVKLGIEETDKFLOGLIDDKRGQQETGTTMDHLLVLQKSDIEYTDQIKGI 299
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 297 ALAMLFGGTDSDSTGTLNLSNLNHNHPEVLKKADELDTQVGDRLNLSNEDLPKLPYLAK 356
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 300 ILIMVIAGNTSVAVTLNLSNLNHNHDPVISKARDEINRVGLDRLEADLSELPLKXN 359
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
```

QY 357 IILETLRLYPAPILPHVSSDITIEGFNIPRDTIIVINGMQRDPOLMNDATCFKPE 416
 Db 360 IVLETLRLHPATLLVPHMASEDCKIGSYDMPRGTTLLVNAWAHEDPNTWDDPDSFKPE 419
 QY 417 RFVVEGEEKLVAFGMGRACPGEPNMAOSVFTLLGLLIQCFDWKRVSEKLDMTENNWI 476
 Db 420 RFEKEBEAQKLLAFGLGRACPGSLAQRTVGLALGSLIQCPEWERVGNVEVDMKEGVGN 479
 QY 477 TLSRLIPLEAMCKARPLATKI 497
 Db 480 TVPKAIPLKAICKARPLFKHI 500

RESULT 15
 AAG42047
 ID AAG42047 standard; Protein; 505 AA.
 XX
 AC AAG42047;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 52390.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 08-MAY-1999; 99US-0132484.
 PR 08-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0134941.
 PR 21-MAY-1999; 99US-0135124.
 PR 24-MAY-1999; 99US-0135353.
 PR 25-MAY-1999; 99US-0135629.
 PR 27-MAY-1999; 99US-0136021.
 PR 28-MAY-1999; 99US-0136392.
 PR 01-JUN-1999; 99US-0136782.
 PR 03-JUN-1999; 99US-0137222.
 PR 05-AUG-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.

Matches	236;	Conservative	100;	Mismatches	150;	Indels	15;	Gaps	7
Qy	4	WVSAYLVFLVFLGVKFVFSQSKLRNLP	PPPPP-LPIIGNLNLLLEQPIHRRFQRM	SKQY-	61				
Db	11	WVLYSI-FSIFLIISFKFLKPKQNLP	PPSPGWLPIIGHLLKLPPIHRTLGRFS	ETLD	69				
Qy	62	-----GNVSVLWFGSRLAVWISS-PTA	VOEFTKIDVALANRPLSLSGKXIYFNNT	TVGSC	116				
Db	70	HNDGGGVMSRLGSLRVVYVSSHVA	AECEFGKNDVVLANRPQVIIGHVGYN	NTNMAA	129				
Qy	117	SHGEHNELRIATLDVLSRTORVHS	PSGISRDETKRLMORLAVLAKNSNEE	EPARVEISSM	176				
Db	130	PYGDHWNRLRLCTIEFSTRLNCF	LYVRIDEVRLISRSRLAGTKK--TVV	VELKPM	186				
Qy	177	FNDLTNNINRMISGKGFYGBESEM	QNVKEAREFRETVMELMLGLANKGD	HLPLRMF	236				
Db	187	LMDLTFNNIMRMWTGKYYGEET--	TDEEAKRVKLVADVANTSSGNAV	DYVPILRL-	243				
Qy	237	DFQNVKRLKSIISKRYDSIINKL	IHNENRASNDRONSMDHLLKLQETO	POYYTDQIIKGL	296				
Db	244	-FSSYENRVKLGEBTKFLQGLD	DDKRGQGETGTTMIDHLLVLQKSD	IEYYTDQIIKI	302				
Qy	297	ALAMLFGTDSSTGTLEWSILN	LHNHEVLKARDELDTVOGDPLN	ESDLPLKPYLRK	356				
Db	303	ILINVIAGTNTSAVTLTSLW	SLNHLNHPDIVISKARDEIDNR	VGLDRLLI	362				
Qy	357	ILLETRLYPAPTLLIPIHVS	EDITIEGFNPRTDIVIINGWMORD	QOLMNDATCFKPE	416				
Db	363	IVLETLRLHPATPLVLVPHMA	SEDCKISYDNPRGTTLLVNWAWA	THRDDNTWDDP	422				
Qy	417	RFDFVEGEBEKKLVAFGM	RRACPGSPFMAMQSVFSLGLLI	QCDFWKRVSEEK	476				
Db	423	RFEKEEBAQKLLAFGLGR	ACPGSLAQRIVGLALGSLIQCFE	WERVGVNEVD	482				
Qy	477	TLSRLIPLFAMCKARPLATKI		497					
Db	483	TVPKAIPLKATCAKRPFLHKI		503					

Search completed: October 20, 2003, 20:31:37
Job time : 90 secs

Query Match 46.2%; Score 1196.5; DB 21; Length 505;
Best Local Similarity 47.1%; Pred. No. 1.9e-99;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:34:14 ; Search time 69 Seconds

(without alignments)
1185.429 Million cell updates/sec

Title: US-09-868-546A-2

Perfect score: 2592

Sequence: 1 LVVVSYAVLFLVFLGVKF.....RLIPLEMCKARPLATKIGI 499

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters: 609560

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	921.5	35.6	514	12	US-10-356-153-68
2	921.5	35.6	514	15	US-10-142-231-68
3	810.5	31.3	509	12	US-10-097-559-11
4	773	29.8	512	12	US-10-356-153-66
5	773	29.8	512	15	US-10-142-231-66
6	756	29.2	506	10	US-09-814-786-65
7	753	29.1	471	12	US-10-097-559-42
8	742	28.6	516	14	US-10-067-534-3
9	730	28.2	523	10	US-09-814-786-67
10	729	28.1	510	12	US-10-021-425-30
11	724.5	28.0	509	12	US-10-097-559-41
12	718	27.7	502	12	US-10-259-165-146
13	711.5	27.4	508	9	US-09-796-256A-2
14	707	27.3	520	14	US-10-067-534-4
15	704	27.2	508	9	US-09-931-267-4

16	702	27.1	508	9	US-09-931-267-5
17	695	26.8	325	12	US-10-259-165-102
18	695	26.8	325	12	US-10-259-165-434
19	678	26.2	542	12	US-10-289-757-107
20	673.5	26.0	510	10	US-09-814-786-66
21	671.5	25.9	477	9	US-09-931-267-6
22	669.5	25.8	477	9	US-09-931-267-7
23	666.5	25.7	543	12	US-10-289-757-183
24	664.5	25.6	532	12	US-10-289-757-106
25	657	25.3	526	12	US-10-361-460-73
26	657	25.3	529	16	US-10-174-693-405
27	652.5	25.2	535	14	US-10-067-534-2
28	651.5	25.1	514	10	US-09-947-027-4
29	651.5	25.1	514	14	US-10-091-009-4
30	845	24.9	519	12	US-10-097-559-9
31	638.5	24.6	531	11	US-09-899-642-2
32	631	24.3	495	12	US-10-021-425-29
33	618	23.8	511	9	US-09-796-256A-4
34	613	23.6	553	10	US-09-349-385-8
35	611	23.6	518	15	US-10-163-198-35
36	609.5	23.5	534	10	US-09-349-385-1
37	609.5	23.5	534	15	US-10-163-198-44
38	603.5	23.3	474	12	US-10-097-559-3
39	603	23.3	473	12	US-10-097-559-1
40	601	23.2	504	15	US-10-163-198-41
41	598	23.1	530	10	US-09-349-385-4
42	598	23.1	530	15	US-10-022-025A-3
43	598	23.1	530	15	US-10-163-198-46
44	596.5	23.0	516	15	US-10-163-198-23
45	580.5	22.4	514	15	US-10-163-198-97

ALIGNMENTS

RESULT 1

US-10-356-153-68
; Sequence 68, Application US/10356153
; Publication No. US20030166176A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/356,153
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US/10/142,231
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-356-153-68

Query Match	35.6%	Score 921.5;	DB 12;	Length 514;
Best Local Similarity	37.9%	Pred. No. 1.6e-71;		
Matches 196;	Conservative 103;	Mismatches 175;	Indels 43;	Gaps 11;
QY	2	LVVVSYAVLFLVFLGVKFVQSRLN-----IPGPPPLPIIGNLLEQ--PIHR	52	
Db	8	VILFTLALLLV-----VIQRRIRRHKLOKVKAPQPSWVFNGLHLLTQKVP	60	
QY	53	FFORMSKYQGNVVMFSGSLRAVWISSPTAYOECFTKHVALANRLPSLS-GK	111	
Db	61	ILSSISESGVPMHVLQGLRPAVLVASSDLAKECFTNDKAFASR-PLRSAGKH	119	
QY	112	TVGCSCSHGHWNRRLRITDALVLSQTVHSFGISDETQRLMQRLVLAKNNEE	171	
Db	120	IFSMAPGYGWNLRKMCCTQILSATRIDSFRHVRVEVSALIRSLF---DSCOR	176	

|||| : : || : : || : :
Db 477 DLDMEIFGLSTPKPLATVIEPR-LSPKL 506

RESULT 4

US-10-356-153-66
; Sequence 66, Application US/10356153
; Publication No. US20030166176A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/356,153
; PRIOR FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US/10/142,231
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Taxus cuspidata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(512)
; OTHER INFORMATION: "Xaa" equals any peptide
US-10-356-153-66.

Query Match 29.8%; Score 773; DB 12; Length 512;
Best Local Similarity 32.5%; Pred. No. 1.3e-58;
Matches 160; Conservative 125; Mismatches 178; Indels 30; Gaps 10;
QY 8 AVLFLVFLGVKVFQSRKLRN----IPGPPPLPIIGNLNLEQPIHRRFFORMSKOYGN 63
Db 22 AVLTILFLGIFYL--RGLRNGRLPPGPPIWPVIGNLHQGLKLPNRNLELAKKHP 79
QY 64 VVSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLGGKYIFYNNTTVGSCSHGEHR 123
Db 80 IMLMGLGSPAVIVSSSAAKEVLKTHDLVFASRPESAAGKYIAYNKDIVFSPYGPYWR 139
QY 124 NLRRITADVLTSTQVHVSFGIRSDETKRLMORLVLAKNSEEEFARVEISSMFDNTYN 183
Db 140 QMKKICVVELLNARRIESLRVREEVSVII-RSVWEK--SKOGAVAVNLSKTLSSLTQ 196
QY 184 NIMRMISGKRFYGESEMKNVBEAREFRTVTMELMGLANKGDHLPFLRWDFQNVK 243
Db 197 LMQIFSSNDGNS-----VTAIKEMMSEVSETAGAFNIGDYFPWMDWMDLQGIOR 249
QY 244 RLKSIKRYDSILNKILHEN--RASNDRO--NSMIDHLLKLOETQPOQYVTDQIIKGLAL 298
Db 250 RMTKAHDYFDQVITKIEHQHTRAMEDTQOPKDIIDALLQMENTDGVTTITMENIKAVVL 309
QY 299 AMFLGGTDSSTGLWSLNLNHPVLKARDELDTQVGDRLNLES DLPKLPYLKII 358
Db 310 GIFLGAETTSITLWAMSAMLENPEVAKVQEEIESVVGKRVKEMWESMEYLQCVV 369
QY 359 LETRLYPAPILIPHVSSEDIIEGFIIPRDTIVINGMQRDPOLMNDATCFKPERF 418
Db 370 KKTMRLYPAVPLIIPHESTQDCTVNGYFIPERTRILVNAWAIGKDPNVWDDALAFKPRF 429
QY 419 ---DVEGEKK----LVAFGMGRACGPEPMQSVFTLGLLIQCFDMKRVSEKLDMT 471
Db 430 LGXNVDLQKGEFFDMVFPFAGKGCPCGASMAVVTMEHALAQLMHCFOWR--IEGELDMS 487
QY 472 ENRWITLSRLIPL 484
Db 488 ERLAASVQKKVDL 500

RESULT 5

US-10-142-231-66

; Sequence 66, Application US/10142231
; Publication No. US20030077796A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/142,231
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Taxus cuspidata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(512)
; OTHER INFORMATION: "Xaa" equals any peptide
US-10-142-231-66

Query Match 29.8%; Score 773; DB 15; Length 512;
Best Local Similarity 32.5%; Pred. No. 1.3e-58;
Matches 160; Conservative 125; Mismatches 178; Indels 30; Gaps 10;
QY 8 AVLFLVFLGVKVFQSRKLRN----IPGPPPLPIIGNLNLEQPIHRRFFORMSKOYGN 63
Db 22 AVLTILFLGIFYL--RGLRNGRLPPGPPIWPVIGNLHQGLKLPNRNLELAKKHP 79
QY 64 VVSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLGGKYIFYNNTTVGSCSHGEHR 123
Db 80 IMLMGLGSPAVIVSSSAAKEVLKTHDLVFASRPESAAGKYIAYNKDIVFSPYGPYWR 139
QY 124 NLRRITADVLTSTQVHVSFGIRSDETKRLMORLVLAKNSEEEFARVEISSMFDNTYN 183
Db 140 QMKKICVVELLNARRIESLRVREEVSVII-RSVWEK--SKOGAVAVNLSKTLSSLTQ 196
QY 184 NIMRMISGKRFYGESEMKNVBEAREFRTVTMELMGLANKGDHLPFLRWDFQNVK 243
Db 197 LMQIFSSNDGNS-----VTAIKEMMSEVSETAGAFNIGDYFPWMDWMDLQGIOR 249
QY 244 RLKSIKRYDSILNKILHEN--RASNDRO--NSMIDHLLKLOETQPOQYVTDQIIKGLAL 298
Db 250 RMTKAHDYFDQVITKIEHQHTRAMEDTQOPKDIIDALLQMENTDGVTTITMENIKAVVL 309
QY 299 AMFLGGTDSSTGLWSLNLNHPVLKARDELDTQVGDRLNLES DLPKLPYLKII 358
Db 310 GIFLGAETTSITLWAMSAMLENPEVAKVQEEIESVVGKRVKEMWESMEYLQCVV 369
QY 359 LETRLYPAPILIPHVSSEDIIEGFIIPRDTIVINGMQRDPOLMNDATCFKPERF 418
Db 370 KKTMRLYPAVPLIIPHESTQDCTVNGYFIPERTRILVNAWAIGKDPNVWDDALAFKPRF 429
QY 419 ---DVEGEKK----LVAFGMGRACGPEPMQSVFTLGLLIQCFDMKRVSEKLDMT 471
Db 430 LGXNVDLQKGEFFDMVFPFAGKGCPCGASMAVVTMEHALAQLMHCFOWR--IEGELDMS 487
QY 472 ENRWITLSRLIPL 484
Db 488 ERLAASVQKKVDL 500

RESULT 6

US-09-814-786-65
; Sequence 65, Application US/09814786
; Patent No. US20020100072A1
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; KIKUCHI, Shigeto
; SHIMADA, Yukihisa
; OHBAYASHI, Masaya
; SHIMADA, Ritsuko

RESULT 8
 US-10-067-534-3
 ; Sequence 3, Application US/10067534
 ; Publication No. US20020187538A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Essenberg, Margaret K.
 ; APPLICANT: Chen, Xiao-Ya
 ; APPLICANT: Luo, Ping
 ; APPLICANT: Wang, Yan-Hong
 ; TITLE OF INVENTION: cDNA Clone of (+)-Delta-Cadinene-8-Hydroxylase Gene from Cotton
 ; FILE REFERENCE: 006602-113
 ; CURRENT APPLICATION NUMBER: US/10/067,534
 ; CURRENT FILING DATE: 2002-02-07
 ; PRIOR APPLICATION NUMBER: US 60/267,160
 ; PRIOR FILING DATE: 2001-02-07
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 516
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-067-534-3

Query Match 28.6%; Score 742; DB 14; Length 516;
 Best Local Similarity 32.3%; Pred. No. 6.6e-56;
 Matches 161; Conservative 109; Mismatches 199; Indels 30; Gaps 8;
 QY 7 YAVLEL-VLFLGVKVFQSRKLNIPPPPLPIGNLNLLEQPIHRRFFORMSKQYGNV 65
 DB 18 YAVILTTVFSIIWYIFKSPQSLPPGRLPIVGNLPLFLDPLDHTYFANLAQSHGP 77
 QY 66 SLWFGRLAVISSPTAYOECFTKHDVALANRLPSLGKIFYNNNTTVGSCSHGEHWRNL 125
 DB 78 KNLGSKLITVNSPSLAREILKQDINFNSRDVPLTGRAATYGGIDIVWTPVGAERQL 137
 QY 126 RRITADVLTSTQRVHSGIRSDTKRLMORLVAKNSNEEFPARVEISSMFDN---LTY 182
 DB 138 KKICVLKLSRKTLDSEFYELRRKXVRETRYLY-----EQGRKQSPVKVGDLFLTM 189
 QY 183 NNI-MRMISGRFYGESEMKNVEEAREPRETETVTEMLMGLANKGDHLPFLRWDFQNV 241
 DB 190 MNTMMLWGGSVKAEEME-----SVGTEFGVISEITRLISEPHVSDFFPWLARFDLQGL 245
 QY 242 ERLKLSIKRYDSILNKILHEN-----ASNDRONSMIDHLLKLOETQPOY---YTDQIIK 294
 DB 246 VKRMGVCARELDVLDRAIQMPLRGCRDDDEVKDFLOVLMKLKQDEGDSVPIITINHVK 305
 QY 295 GLALAMLPFGTSDSTGTLEWSLNLNHPVLKARDELDTVQODRLNESDLPKLPYL 354
 DB 306 ALITDMVVGTDSTNTIEFAMALMSNPILIKRAQOEELDEVVVGKDNIVEESHITLPIYI 365
 QY 355 RKIIETRLYPAPILIPHVSSDITIEGFNIPRDTIVINGWGORPOLMNDATCFK 414
 DB 366 LAIMKETLRHPTLPILVPHRPANTVVGTYTIPKTKIFVNVWSIQORDNVWENTFEPR 425
 QY 415 PERF-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCQFDMKRVSEKL 468
 DB 426 PERFDNNSCDFTGANYSYFPFGSGRRICAGVALAERWVLYTLATLLHSPDKWIPGSHVL 485
 QY 469 DMTENNWITLSRLIPLEAM 487
 DB 486 DLXKEFTGIVLKIPLVAL 504

RESULT 9
 US-09-814-786-67
 ; Sequence 67, Application US/09814786
 ; Patent No. US20020100072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIKUCHI, Yasuhiro
 ; KIKUCHI, Shigeto
 ; SHIMADA, Yukihisa

OHYAYASHI, Masaya
 SHIMADA, Ritsuko
 OKINAKA, Yasushi
 TITLE OF INVENTION: NOVEL PLANT GENES
 NUMBER OF SEQUENCES: 67
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10112-3801
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS Ver3.30
 SOFTWARE: PATENT AID Ver1.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/814,786
 FILING DATE: 23-Mar-2001
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/616,990
 FILING DATE: 14-Jul-2000
 APPLICATION NUMBER: JP44963/92
 FILING DATE: 02-MAR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Perry, Lawrence S.
 REGISTRATION NUMBER: 31865
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-218-2100
 TELEFAX: 212-218-2200
 INFORMATION FOR SEQ ID NO: 67:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 523 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Campanula medium
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 180 to 1748
 IDENTIFICATION METHOD: by experiment
 SEQUENCE DESCRIPTION: SEQ ID NO: 67
 US-09-814-786-67
 Query Match 28.2%; Score 730; DB 10; Length 523;
 Best Local Similarity 31.3%; Pred. No. 7.5e-55;
 Matches 159; Conservative 117; Mismatches 208; Indels 24; Gaps 7;
 QY 4 VVSVAVLFLVLGVKVFQSRKLNIPPPPLPIGNLNLLEQPIHRRFFORMSKQYGN 63
 DB 12 LVAATSLVLTATYSFIRFLKPSHHHLLPPGTPGWIIGALPLIGTMPHVLADMAVKYGP 71
 QY 64 VVSLWFGRLAVISSPTAYOECFTKHDVALANRLPSLGKIFYNNNTTVGSCSHGEHWR 123
 DB 72 INYLKLGSKGTVWASNPAAAFKTHDANFSNRPIDGGPTTYLAYNAQDMVFAEYGPWK 131
 QY 124 NLRRITADVLTSTQRVHSGIRSDTKRLMORLVAKNSNEEFPARVEISSMFDNLTYN 183
 DB 132 LLRLKCLSLHMLGPKALEDAWHVKVSEVGHMLKEMK---EQQSKSVFVPPVPEMLTYMAN 189
 QY 184 NIMRMISGRFYGESEMKNVEEA---REFRETETVTEMLMGLANKGDHLPFLRWDFQ 239
 DB 190 MIGRIILSRPPIVTSKLDSSASASVSEFQYMWMLMRMAGLNFIGNDFIVIAWMDIQ 249
 QY 240 NVEKRLKISKRYDSILNKILHEN-RASNDRONSMIDHLLKLOETQPOYTTQI----- 292
 DB 250 GTQDMKVIQKKFDVLLNKMIKEHTESAHRKDN---PDFLDILMAATQENTEGIQNLVYN 307
 QY 293 IKGLALAMLPFGTSDSTGTLEWSLNLNHPVLKARDELDTVQODRLNESDLPKLP 352

```
Db 308 VKALLDLFTAGTTSSTSVIEWALAEMLNHRQILNRAHEEMDQVIGNRRLEQSDIPNLP 367
Qy 353 YLRKIILETLRLYPAPILPHVSSSEDTTIEGNIIPRDTIVINGMGMORDPOLWMDATC 412
Db 368 YFOAIKCTFKHSTPLNLPRISTEACEVDGPHIPKNTRLIVNIAIGDRPKWENPLD 427
Qy 413 FKPERP-----DVEGEKKLVAFGMRACCPGEPMAQSVFTLGLLIQCFDKRV - 463
Db 428 FTERFLSEKHAIDPRGNHPELIPFGAGRRICAGARMGAASVEYILGLTVHFSDFWKLPD 487
Qy 464 SEEKLDMTENNWTLSRLIPLEAMCKAR 491
Db 488 GWVEVNMEEFGIALQKKVPLSAIVTPR 515

RESULT 10
US-10-021-425-30
; Sequence 30, Application US/10021425
; Publication No. US20030148420A1
; GENERAL INFORMATION:
; APPLICANT: Suzanne L. Bolten
; APPLICANT: Alan M. Easton
; APPLICANT: Leslie C. Engel
; APPLICANT: Dean M. Messing
; APPLICANT: John S. Ng
; APPLICANT: Beverly A. Reitz
; APPLICANT: Scott A. Vaccaro
; APPLICANT: Mark C. Walker
; APPLICANT: Ping T. Wang
; APPLICANT: Robin A. Weinberg..
; TITLE OF INVENTION: Aspergillus ochraceus 11 alpha
; FILE REFERENCE: S03196-00-US
; CURRENT APPLICATION NUMBER: US/10/021,425
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: USSN 60/244,300
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max AAB94588
US-10-021-425-30

Query Match 28.1%; Score 729; DB 12; Length 510;
Best Local Similarity 31.1%; Pred. No. 8.8e-55;
Matches 158; Conservative 116; Mismatches 192; Indels 42; Gaps 10;

Qy 7 YAVLFLVLFLGVKVF-----QSRKLRNIPPPPLPIIGNLNL--EQPIHFFQ 55
Db 11 PSIVFITSILFIFFVFKLVQSDSKTSSTCKLPPGPRTLPLIGNIHQVIGSLPVHYLK 70
Qy 56 RMSQYGNVSLWFGSLAVISSPTAYQECFTKHDVALANRLPSLSGKIYFYNTTVGS 115
Db 71 NLADYQKLEMLKLGVEVNIITVPEMAQEIIMTKHDNLFSDRDPVLSRVYNGSGIVF 130
Qy 116 CSHEGHRNLRITDVLSTQVHSGFSGISDETKLMORLVLAKNSNEEFARVEISS 175
Db 131 SOHGDYWLQAKICTVELLTAKRVQSFRIEEVAELVKK--IAATASEGGSFNLQ 188
Qy 176 MFNDLYNNIMRISGKRFYGESEMKVVEARFRETVTMELMGLANKDGLPFLRW 235
Db 189 SIYSMTFG-----IAARAFAFKKRYQQV-----FISNMHQLMLGGFVADLPSSRV 238
Qy 236 FDFQNVKRLKSKRYDSILNKLHE-----NRASNDQ--NSMIDHLKLQETQPYT 289
Db 239 FOMGATGKLEKXHVTRDRIQDIIDEHKNRNSSEEREAVEDLDVLLRFQKSEFRLT 298
Qy 230 DQIINGLALAMFGGTDSTGTLEWSLNLNHPVLKARDELDTQVQDRLNESDLP 349
Db 299 DDNIKAVIQDIFIGGETSSSVVWGMSELINPRVMEEAQAEVRVYDSKGYVDETELH 358
```

```
Qy 350 KLPYLRKIILETLRLYPAPILPHVSSSEDTTIEGNIIPRDTIVINGMGMORDPOLWMD 409
Db 359 QLIYLKLSIIKETRLWLPVPLVPRVSRERCOINGVEIPSKTRIINAWAIGRNPKYWGE 418
Qy 410 ATCFKPERP-----DVEGEKKLVAFGMRACCPGEPMAQSVFTLGLLIQCFDWK--- 461
Db 419 TESFKPERFLNNSIDFRGTDFEPIFGAGRRICPGITFAIPNIELPLAQLLYHFDWKLPN 478
Qy 462 RVSEKKLDMTENNWTLSR-----LIPL 484
Db 479 KMWNEHLDMTESNGITLRRQNDCLIP 506

RESULT 11
US-10-097-559-41
; Sequence 41, Application US/10097559
; Publication No. US2003016255A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: Ralston, Lyle P.
; TITLE OF INVENTION: Cytochrome P450s and Uses Thereof
; FILE REFERENCE: 07678/100003
; CURRENT APPLICATION NUMBER: US/10/097,559
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 60/274,241
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/275,597
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Nepeta racemosa
US-10-097-559-41

Query Match 28.0%; Score 724.5; DB 12; Length 509;
Best Local Similarity 32.4%; Pred. No. 2.2e-54;
Matches 164; Conservative 101; Mismatches 208; Indels 33; Gaps 10;

Qy 4 VVSYAVLVLVLFLGVKVFQSRKLRNI-----PGGPPPLPIIGNLNLLEQPIHFFQ 57
Db 1 MVSLSYFLIALICTLTPFLFLNKRNRYSKGKTPPPPKLPVIGNLHQGLYPHRYLQSL 60
Qy 58 SKQYGNVSLWFGSLAVISSPTAYQECFTKHDVALANRLPSLS-GKIYFYNTTVGSC 116
Db 61 SRYGELMQLHFGSVPLVASSPEAREIMKNQDIVFSNR-PKMSIANRLFENNRDVAFT 119
Qy 117 SHGHEWRNLRITDVLSTQVHSGFSGIRSDTKLMORLVLAKNSNEEFARVEISS 176
Db 120 QYGEYWRQIRSIQVLLQSNKRVSPPRVREETSIMVEKIMQLGSSSS---TPVNLSEL 176
Qy 177 FNDLTNNIMRISGKRFYGESEMKVVEARFRETVTMELMGLANKDGLPFLRW 236
Db 177 LLSLTNDVVCVRVLGKKGKGGNGS-----BEVDKLEMLTEIQNLMGISPWEIFPWNWT 232
Qy 237 -DFQNVKRLKSKRYDSILNKLHE-----ASNDQRNSMIDHLKLQETQPY - 287
Db 233 RFDGVDQVDRIVKAFDGFLESVIQEKERDGDGDGDGALDFVDILLQFORENQRS 292
Qy 288 -YTDQIIKGLALAMFGGTDSTGTLEWSLNLNHPVLKARDELDTQVQDRLNES 346
Db 293 PVEDDTVKALILDPMFVAGTDTTATALEWAEVLKRNPRAMKRLQNEVREVAGSKAIEEE 352
Qy 347 DLPLFLVKRIILETLRLYPAPILPHVSSSEDTTIEGNIIPRDTIVINGMGMORDPOL 406
Db 353 DLEKMYLKASIKESURLHVPVLLVPRESTRTNVLGYDIASGTVLINAWAIARDPSV 412
Qy 407 WNDATCFKPERF-----DVEGEKKLVAFGMRACCPGEPMAQSVFTLGLLIQCFDW - 460
Db 413 WENPEFLPERFLDSSIDYKGLHFELLPPFGAGRGCGPGATFAVIDELALAKLVHFDG 472
Qy 461 --KRVSEKKLDMTENNWTLSRLIPL 484
```


US-10-067-534-4
; Sequence 4, Application US/10067534
; Publication No. US20020187538A1
; GENERAL INFORMATION:
; APPLICANT: Essenberg, Margaret K.
; APPLICANT: Chen, Xiao-Ya
; APPLICANT: Luo, Ping
; APPLICANT: Wang, Yan-Hong
; TITLE OF INVENTION: cDNA Clone of (+)-Delta-Cadinene-8-Hydroxylase Gene from Cotton
; CURRENT APPLICATION NUMBER: US/10/067,534
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/267,160
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-067-534-4

Query Match 27.3%; Score 707; DB 14; Length 520;
Best Local Similarity 31.2%; Pred. No. 7.4e-53;
Matches 159; Conservative 108; Mismatches 204; Indels 38; Gaps 10;

Qy 2 LVVVSAYAVLFLVFLGVK-FVFSQKRLRNIPPGPPPLPIIGNLNLLEQPIHRFFORMSKQ 60
Db 15 LTPYAVAVLITATFSILWYIFKSPQPLPPGRGLPIVGNLFLPDPLHTYETKLAQS 74

Qy 61 YGNVVSFWGSLRLAVVSSPTAYQECFTKHDVALANRLPSLSGKIYFYNNTTVGSCSHGE 120
Db 75 HGPIFKINGSLKLVVWNSPSSLSEILKQDINFSDVPLTARAVTYGGLDVLWLPYGA 134

Qy 121 HWRNLRRITADVLTSTQVHSFGIRSDETK--RLMORLVAK--NSNEEEFARVEIS 174
Db 135 EWRMLRVCAAKLFSRKTLDSTFYELRRKEIRERTRCLYOKGLEKSPVNVGSQLF----- 188

Qy 175 SMFNDLTNNI--MRMISGRFYGESEEMKNVEAREFRETVTMELEMLGLANKGDHLPFL 233
Db 189 -----LTMNMLMNLWGSVKADME---SVGTGFKGVISEITRLLGVENVSDFFPML 239

Qy 234 RWFDFONVEKLKLSIKRYDSILNKILHE-----NRASNDQ--NSMDHLLKLOETOPQY 287
Db 240 ARFDLQGLVKKVHLYARDLDAILDRAIEQMORLSRDGDDGCECKDFLQHLMKLRDQADS 299

Qy 288 --YTDQIIKGLALAMLFGGTDSSTGTLEWSLSNLLNHPVLKKADELDTQVGDRLN 344
Db 300 DVPITMNVKAVLMDVMVGGTSTNTIEFVMAELISNPELMRAQDELDEVGKDIVE 359

Qy 345 ESDLPLPYLRKILLETLLRYPAPILIPHVSSDITIEGFNIPRTIIVINGWGMORDP 404
Db 360 ESHITSLPYLAVLAKSTLRLYPTIPLVPHRSETALVGGVTIPKNTKIFINVWSIQRP 419

Qy 405 QLWNDATCFKPERF-----DVEGEKKLVAFGMRRACPGEPMAQSVFTLGLLQCF 458
Db 420 NVWEYPTFPERFPLDKKSCDFTGTDYSYLPFGSGRRICAGIALAERMILYTLTLHSF 479

Qy 459 DWKRVSEKLDMTENNNTLSRLIPLEAM 487
Db 480 DWTIPDGHVLDLEKFGIVLKLTKPLVAL 508

RESULT 15
US-09-931-267-4
; Sequence 4, Application US/09931267
; Patent No. US20020062496A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clinton
; APPLICANT: Franke, Rochus
; APPLICANT: Ruegger, Maxwell
; TITLE OF INVENTION: Genes Encoding P-Coumarate 3-Hydroxylase (C3H) and Methods of Use
; FILE REFERENCE: N1422-005

; CURRENT APPLICATION NUMBER: US/09/931,267
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/225554
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-931-267-4

Query Match 27.2%; Score 704; DB 9; Length 508;
Best Local Similarity 30.9%; Pred. No. 1.3e-52;
Matches 156; Conservative 113; Mismatches 202; Indels 34; Gaps 9;

Qy 1 LVVVSAYAVLFLVFLGVKVFQSKLR-NIPGPPPLPIIGNLNLLEQPIHRFFORMSK 59
Db 11 IAAVWSYKLI-----QRLRYKFFGSPSPKPIVGNLYDIKPVRFRCYEWAAQ 56

Qy 60 QYGNVVSFWGSLRLAVVSSPTAYQECFTKHDVALANRLPSLSGKIYFYNNTTVGSCSHG 119
Db 57 SYGPIISVMIGSLNVVSSAELAKEVLEKHDOKLADRHRNRSTFAFSRNGQDLIWADYG 116

Qy 120 EWRNLRRITADVLTSTQVHSFGIRSDETKRLMORLVAKNSNEEEFARVEISSMFD 179
Db 117 PHYVKRVKVTLELFTPKLESIRPTREDEVAMVESVFRDCNLPENRAKGLQLRKYLGA 176

Qy 180 LTVNNIMRMISGRFYGESEEMKNVEAREFRETVTMELEMLGLANKGDHLPFLRWFDFQ 239
Db 177 VAFNNITRIAFGRFMAEGVVD--EGLFKAIVSNGKLKGLASLSIAEHIPWRWM-FP 233

Qy 240 NVEKRLKSISKRYDSILNKILHEN-----RASNDQNSMIDHLLKLOETOPQY-YTDQIIK 294
Db 234 ADEKFAEAGARRDLTRAMEEHTLAROKSSGAKQHFVDALITLKD---QYDLS EDTII 290

Qy 295 GLALAMLFGGTDSSTGTLEWSLSNLLNHPVLKKADELDTQVGDRLNLSLPLPYL 354
Db 291 GLLWDMITAGMDTTAITAEWAMEIKNPRVQKQVEEFDRVVGLDRILTEADFSRLPYL 350

Qy 355 RKIILETLRLYPPAPILIPHVSSDITIEGFNIPRTIIVINGWGMORDPOLWNDATCFK 414
Db 351 QCVVKSFRLLHPTPLMLPHRSNADVKIGYDIPKGSNVHVNVWAVARDPAVWKNPEFR 410

Qy 415 PERF-----DVEGEKKLVAFGMRRACPGEPMAQSVFTLGLLQCFDW---KRVSEE 466
Db 411 PERFLEEDVMKGDHDFLLPFGAGRRVCFGAQLGVLNLTSMMSHLLHFWTTPPQGTKE 470

Qy 467 KLDMTENNNTLSRLIPLEAMCKAR 491
Db 471 EIDMSENPGLVTVYMRTPVOAVTPR 495

Search completed: October 20, 2003, 20:44:04
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:30:09 ; Search time 30 Seconds
(without alignments)
703.771 Million cell updates/sec

Title: US-09-868-546A-2
Perfect score: 2592
Sequence: 1 LLVVSVAVLFLVFLGVKF.....RLIPLEAMCKARPLATKIGI 499

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778.5	30.0	513	4	US-09-351-229-4
2	764.5	29.5	532	3	US-08-948-564-10
3	756	29.2	506	1	US-08-313-075A-38
4	756	29.2	506	3	US-08-606-505B-65
5	756	29.2	506	3	US-09-616-990-65
6	747	28.8	471	4	US-09-126-420A-20
7	731.5	28.2	508	4	US-09-126-420A-19
8	730	28.2	523	3	US-08-606-505B-67
9	730	28.2	523	3	US-09-616-990-67
10	729	28.1	510	3	US-08-948-564-4
11	725.5	28.0	495	1	US-08-532-065B-2
12	713	27.5	490	4	US-09-126-420A-16
13	711.5	27.4	508	3	US-08-991-677-2
14	684.5	26.4	509	3	US-08-948-564-18
15	681.5	26.3	496	1	US-08-313-075A-50
16	673.5	26.0	510	3	US-08-606-505B-66
17	673.5	26.0	510	3	US-09-616-990-66
18	665.5	25.7	513	3	US-08-948-564-2
19	664	25.6	504	4	US-09-126-420A-18
20	662.5	25.6	513	4	US-09-564-808-6
21	660	25.5	520	2	US-09-091-432-2
22	660	25.5	520	2	US-09-387-663-2
23	657	25.3	529	4	US-09-615-192A-405
24	648	25.0	520	4	US-09-564-808-2
25	648	25.0	520	4	US-09-564-808-4
26	638.5	24.6	531	4	US-09-380-420C-2
27	633	24.4	516	3	US-08-948-564-12

28	621	24.0	504	4	US-09-499-302A-9	Sequence 9, Appli
29	618	23.8	511	3	US-08-991-677-4	Sequence 4, Appli
30	607	23.4	521	3	US-08-948-564-14	Sequence 14, Appl
31	598.5	23.1	509	4	US-09-499-302A-10	Sequence 10, Appl
32	597.5	23.1	502	4	US-09-499-302A-7	Sequence 7, Appli
33	591.5	22.8	476	1	US-08-313-075A-30	Sequence 30, Appl
34	584	22.5	502	4	US-09-499-302A-2	Sequence 2, Appli
35	583	22.5	504	4	US-09-499-302A-6	Sequence 6, Appli
36	578.5	22.3	496	3	US-09-292-768-64	Sequence 64, Appl
37	577.5	22.3	496	3	US-09-292-768-2	Sequence 2, Appli
38	577.5	22.3	496	3	US-09-292-768-66	Sequence 66, Appl
39	577.5	22.3	500	3	US-09-292-768-68	Sequence 68, Appl
40	576.5	22.2	500	3	US-09-292-768-4	Sequence 4, Appli
41	575.5	22.2	500	3	US-09-292-768-70	Sequence 70, Appl
42	574.5	22.2	509	4	US-09-499-302A-8	Sequence 8, Appli
43	570.5	22.0	498	3	US-09-292-768-6	Sequence 6, Appli
44	567.5	21.9	496	3	US-08-881-784-1	Sequence 1, Appli
45	564	21.8	523	3	US-08-948-564-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-351-229-4					
; Sequence 4, Application US/09351229					
; Patent No. 6380465					
; GENERAL INFORMATION:					
; APPLICANT: Barrett, Michael					
; TITLE OF INVENTION: Cytochrome P450 Enzymes and Related Compounds and					
; FILE REFERENCE: P-1049					
; CURRENT APPLICATION NUMBER: US/09/351,229					
; CURRENT FILING DATE: 1999-07-11					
; EARLIER APPLICATION NUMBER: 60/092,596					
; EARLIER FILING DATE: 1998-07-12					
; NUMBER OF SEQ ID NOS: 4					
; SOFTWARE: Patentin Ver. 2.0					
; SEQ ID NO 4					
; LENGTH: 513					
; TYPE: PRT					
; ORGANISM: Zea mays					
US-09-351-229-4					
Query Match 30.0%; Score 778.5; DB 4; Length 513;					
Best Local Similarity 35.3%; Pred. No. 9.3e-64;					
Matches 182; Conservative 99; Mismatches 181; Indels 53; Gaps 15;					
Qy	15	FLGVK-----PVFQSRK-LRNIPGPPPLPIIGNLNLLLEQPIHRRFFORMSKOYGNVVS	66		
Db	5	FLGIALCAAAALFLRGRPPVYVPPPGPKPWPPIIGNLNLMGELPHRSMNELSKRYGPDQM	64		
Qy	67	LWFGSLAVVISPTAYQECFTK-HDVALANRFLPSLSGKIYFYNNTTVSGSCSHGEHW---	122		
Db	65	LWFGS-LPVVVGASAEAKLFLKTNDAAFSRFAVGKYTAYD-----CS-GLLWAPF	116		
Qy	123	-----RMLRITALDVLSSTORVHVSFGIRSDETKRLMORVLAKNSNEEEARVEISSMFN	178		
Db	117	EPYLRQARRICATELFSATRLSEFHIRDEVRMLRQLRQAAAGT-----VLRLDYQL	170		
Qy	179	DLTYNNIMRMSIKRFFGESEMKVNEEA-----REFRETVMLEMLGLANKGDHLPLFL	233		
Db	171	MLALGVISRIVLGKKYVMEEAADGEGSDAPAITPAEFREVMDFEALHGFNIGDYIPWL	230		
Qy	234	RWDFQNVKRLKLSISKRYDSILNKIL--HENRASNDRON-----SMIDHLLKLOE---TQ	284		
Db	231	DWLDLQGYVARMKMKARFCRFLERVLDVNERRREGGNFVAKMDLVLQLADDTLSLE	290		
Qy	285	PQYTTQITIKGLALAMFGGTDSTGTLEWSLNLNHPEVLKARDELDTQGVQDRLLN	344		
Db	291	VQLSRDN-VKAITQDILLIAGTDSNANTLEWAVSELLKNPKILAKAMEELNHRVPRDLVT	349		
Qy	345	ESDLPKLPYLKRIETLRLYPPAPILIPIVSVSSEDITIEGFNIPRODTIVINGMGQRDP	404		

[illegible]

RESULT 2

```

US-08-948-564-10
; Sequence 10, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Cordin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-948-564-10

```

Query Match		29.5%	Score 764.5;	DB 3;	Length 532;
Best Local Similarity		33.8%;	Pred. No. 2e-62;		
Matches 175; Conservative 111; Mismatches 209; Indels 23; Gaps 9;					
<hr/>					
Qy	2	LVVSVAVLFVLVFLGVKFQSR-KLRNIPPGGPPPIIGNLNLL--EQIHRFFFORMS	58		
Dd	13	LISILLAMLVGLYLGKRTSHSGHKICSAQAAGANPILGHHLFGQHHTLTGIMA	72		
<hr/>					
Qy	59	KQYGNNVSLWFGSPALAVVISPTAYQCFTKHUVALANRLPSLSGKIIFYNNTTVGSCH	118		
Dd	73	EKGIPIFTIKLGSYKVIVLSSWENAKECFTVHDKAFSTRPCVAASKLMGYNYAMFGETPY	132		
<hr/>					
Qy	119	GEWRNRURRITALDVLSTQRVHSFSGRSDTKELMORL--VLAKNSNEEFARVEISSM	176		
Dd	133	GPywREIRkUTLIQLLNHRLELLKNTTSESEVAIRELYKLWSRGCPKGVLDVMKWQ	192		
<hr/>					
Qy	177	FNDLTYYNIWMISGKRFPYGBESEMKNVVEAREFRETVTTEMLMWLANKDHPFLRW	236		
Dd	193	FGDLTHNVILVRMKEPYDGASDDYAEGRRARYKKYMGECVSIFGVFLSDAIPFLCWL	252		

[illegible]

RESULT 3

```

US-08-313-075A-38
; Sequence 38, Application US/08313075A
; Patent No. 5639870
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwina C.
; APPLICANT: Tanaka, Yoshikazu
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,075A
; FILING DATE: 30-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 1538/92
; FILING DATE: 27-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL 6698/93
; FILING DATE: 07-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU93/00127
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9433
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-313-075A-38

```


[illegible]

RESULT 4
US-08-606-505B-65
; Sequence 65, Application US/08606505B
; Patent No. 6114601
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: KIYOKAWA, Shigeto
; APPLICANT: SHIMADA, Yukihiisa
; APPLICANT: OHBAYASHI, Masaya
; APPLICANT: SHIMADA, Ritsuko
; APPLICANT: OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
; COMPUTER: IBM PS/V
; OPERATING SYSTEM: MS-DOS Ver3.30
; SOFTWARE: PATENT AID Ver1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,505B
; FILING DATE: 23-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Perry, Lawrence S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 65 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Petunia hybrida
; STRAIN: Falcon Blue
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116 to 1633
; IDENTIFICATION METHOD: by experiment
US-08-606-505B-65

Query Match          29.2%; Score 756; DB 3; Length 506;
Best Local Similarity 32.1%; Pred. No. 1.le-61;
Matches 159; Conservative 114; Mismatches 200; Indels 22; Gaps 51

Qy    10 LFLVLFGLGVVFQSRKLRNIPPPPLPIIGNLINLLEOPIHRPFQRMSSKOYGNVVSLWF 69
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    13 IFLAIIITISLTISKTTGRHLPGRGWPVIGALPLLGAHPHVSLSAKMAKKYGAIMVLKV 72

Qy    70 GSRLAVVISPTAYQCFTKHVDVALANRPLSLGKYIFYNTVTVGSCSHGHWNLRBIT 129
      ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    73 GTCGMAVSTPDAKAFLKTLDINFNSRRPNAGATHLAYNAQDMVFPAHYGPRWKLRLKLS 132

Qy    130 ALDVLSTQRVHSFGSIRSDETKLMRLVLAKXSNBEERFAVEISSMFNDLTYNNIMRMI 189
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    133 NLHWLGKALENWANVRANELGHMLKSM---SDMSREGQRVVAEMLTFAMANNIQQVM 188

Qy    190 SGKFYGESEMKNVBARFPRETVTTEMLGMLANGLKDHLPEFRWDFQNVEKRLLKSIS 249
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    189 LSRRKVFVDKG----VEVNFEFKDMVVELMTIAGFYFNIGDPICLAWMDLQIEIKRMKRLH 243

Qy    250 KRYDSLILKNILHENRASDNDRQNSMIDHLCLKQE-----TOPYYYTDQIIKGLALAMLFCGT 305
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    244 KKFDALLTKMPDEHKATYTTERKGPDPFLDVMMNGMDNSEGERISTTNIKALLMLFTAGT 303

Qy    306 DSSTGTTLWSLSNLIAPHYLKKARDELDTQVGQDRLLNESDLPKLPYLKRKIILETRLY 365
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    304 DTSSSAIEWALAEMWKNPAILKKAQAEMDQVIGNRRLLLESDIENLPYLRAICKETFRKH 363

Qy    366 PPAFILPHVYSSEDITEEGNIIPRDITIIVINGQMORDPOLWINDATCFKPERF----- 418
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    364 PSTPLNIFRSNPFPCIVDGYYIIPKNTKLSVINWAIGRPDWENPLFPNPERFUSGRNSK 423

Qy    419 -DVEGESKKLIAVFGWRRCAPCEGPMMAQSVSFTLGLLIQCFFDKRWKEE-KLMDTENWI 476
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db    424 IDRGNDFELIPFGAGRRIcAGTRMGVIWVVEYIILGTVHSPDWKLPSEVIELNMEEAFGL 483

Qy    477 TLSRLIPLEAMCKAR 491
      |::|::|::|::|
Db    484 ALQAKVPLEAMVTPR 498
      |::|::|::|::|

RESULT 5
US-09-616-990-65
; Sequence 65, Application US/09616990
; Patent No. 6232109
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; INVENTOR: KIKUCHI, Yasuhiro
; SHIMADA, Yukihisa
; OHBAYASHI, Masaya
; SHIMADA, Ritsuko
```

```

/ OKINAKA, Yasushi
/ TITLE OF INVENTION: NOVEL PLANT GENES
/ NUMBER OF SEQUENCES: 67
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
/ STREET: 30 Rockefeller Plaza
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10112-3801
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
/ COMPUTER: IBM PS/V
/ OPERATING SYSTEM: MS-DOS Ver3.30
/ SOFTWARE: PATENT AID Ver1.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/616,990
/ FILING DATE: 14-Jul-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP44963/92
/ FILING DATE: 02-MAR-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Perry, Lawrence S.
/ REGISTRATION NUMBER: 31865
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-218-2100
/ TELEFAX: 212-218-2200
/ INFORMATION FOR SEQ ID NO: 65 :
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 506 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Petunia hybrida
/ STRAIN: Falcon Blue
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 116 to 1633
/ IDENTIFICATION METHOD: by experiment
/ SEQUENCE DESCRIPTION: SEQ ID NO: 65
/
/ US-09-616-990-65
/
/ Query Match 29.2%; Score 756; DB 3; Length 506;
/ Best Local Similarity 32.1%; Pred. No. 1.1e-61;
/ Matches 159; Conservative 114; Mismatches 200; Indels 22; Gaps 5;
/
/ QY 10 LFLVLFLGVKVFQSRKLRNIPGPPPLPIIGNLNLLLEOPIHRRFFORMSKOYGNVSLWP 69
/ DB 13 IFLAIIIIISLTKTGRIHLPGRGWPVIGALPLLGMHPVSLAKWAKKYGAIMYLKV 72
/
/ QY 70 GSRLAVISSPTAYQECFTKHDVALANRLPSLSGKIYFYNNTTVSGSHGHEWNLRRIT 129
/ DB 73 GTCMAVASTDPAKAFKLTIDINSPNPPNAGATHLAYNQDVFAHYGRPKLLKLS 132
/
/ QY 130 ALDVLSTQVHRSFGIRSDTKRLMORLVLAKNSEEFARVEISSMFNDUTYNNIMRMI 189
/ DB 133 NLHMLGGKALENWANVRANELGHMLKSM----SDMSREGQVVAEMLTFFAMANNIGQVM 188
/
/ QY 190 SKRFYGESEMKVVEAREPREVTTEMLELMLGLANKGDHLPFLWPFQNVKRLKSLIS 249
/ DB 189 LSKRFVFDKG-----VEYNEFKDMVVELMTIAGYFNIGDFIPCLAWMDLOGIEKRMKRLH 243
/
/ QY 250 KRYDSILNKILHENRASNDRONSMIDHLKLQE-----TOPOVYTDQIIKGLALAMLFQGT 305
/ DB 244 KFDALLTWFDEHKAITYERKGPFDLDVWNGDNGSEGLSTTNIKALLNLFTAGT 303
/
/ QY 306 DSSTCTLEWSNLNLNHPVLKARDELDTQVGDQRLNLSDELPLKPYLRKIILETLRLY 365
/ DB 304 DTSSSAIBEWALAEKMNPAIKKAQAEQDVIGRNRRLLESIDPNLPYLRAICKETPKKH 363
/
/ QY 366 PPAPLIIHVHSESDITIEGFNIPRTIIVINGWGMQRPQLWNDATCFKPERF----- 418
/
/
/ DB 364 PSTPLNLPRIISNEPCIVDGYIPIKTRLSVNIWAIGRDPQVWENPLEFNPFLSRNSK 423
/
/ QY 419 -DVEGEEKKLVAFGMRACPGEPMAQSVSFTLLGLLQCFDQKRVSE- KLDMTENNWI 476
/
/ DB 424 IDPRGNDPELIFFGARRICATGRMGIVMVEILGTLVHVSFDKLPSEVIELNMEAFGL 483
/
/ QY 477 TLSRLIPLEAMCKAR 491
/ DB 484 ALQKAVPLEAMVTPR 498
/
/
/ RESULT 6
/ US-09-126-420A-20
/ Sequence 20, Application US/09126420A
/ Patent No. 6376753
/ GENERAL INFORMATION:
/ APPLICANT: BATARD, YANNICK
/ APPLICANT: ROBINEAU, TIBURCE
/ APPLICANT: DURST, FRANCIS
/ APPLICANT: WERCK-REICHART, DANIELE
/ APPLICANT: DIDIERJEAN, LUC
/ TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS
/ TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN
/ TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
/ TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S
/ FILE REFERENCE: 03715.0032
/ CURRENT APPLICATION NUMBER: US/09/126,420A
/ PRIOR FILING DATE: 1998-07-30
/ PRIOR APPLICATION NUMBER: 60/054,351
/ PRIOR FILING DATE: 1997-07-31
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: Patent in Ver. 2.1
/ SEQ ID NO 20
/ LENGTH: 471
/ TYPE: PRT
/ ORGANISM: Persea americana
/
/ US-09-126-420A-20
/
/ Query Match 28.8%; Score 747; DB 4; Length 471;
/ Best Local Similarity 34.7%; Pred. No. 6.9e-61;
/ Matches 165; Conservative 98; Mismatches 181; Indels 32; Gaps 8;
/
/ QY 8 AVLFLVLFLGVKVFQSRKLR-----NIPGPPPLPIIGNLNLLLEOPIHRRFFORMSKQ 60
/ DB 2 AILVSLFLAIALTFLLKLNKREKKPNLPPSPNLPPIIGNLHQLGNLPHRSLRLSLANE 61
/
/ QY 61 YGNVSVISWFGSRLAVISSPTAYQECFTKHDVALANRLPSLSGKIYFYNNTTVSGSCHGE 120
/ DB 62 LGPLIILHLGHIPTLIVSTAETAEILKTHDLIFASRPSTTAARRIFYDCTDVAFSPYGE 121
/
/ QY 121 HWENLRRTALDVLSTQVHRSFGIRSDTKRLMORLVLAKNSEEFARVEISSMFNDL 180
/ DB 122 YMRQVKICVLELSIKRVNSYRSIRREEVGLMWERISQSCSTGE----AVNLSELLLL 177
/
/ QY 181 TYNNIMRMISGKFPYGESEMKVVEAREPREVTTEMLELMLGLANKGDHLPFLWPF-FQ 239
/ DB 178 SSGTITRVAFKKYGESE-ERKN-----KPADIALETLLTMGAFFVDYFPFAWVDVLT 231
/
/ QY 240 NVEKRLKLSIKRYSILNKILHE-----NRASNDRONSMIDHLKLQETQP--OYSTD 290
/ DB 232 GMDARLKRNGELDAFVDHVIDDLHLSRKANGSDGVEQKDLVDVLLHLQKSSLGVLHNR 291
/
/ QY 291 QIIKGLALAMLFQGTDSSTGTLEWSNLNLNHPVLKARDELDTQVGDQRLNLSDELPLK 350
/ DB 292 NNLKAVILDMFSGTDTTAVILEWAMAELIKHPDVMEKAQQEVRVVGKAKVVEEDLHQ 351
/
/ QY 351 LPVLRKIILTLRLYPPAPLIIHVHSESDITIEGFNIPRTIIVINGWGMQRPQLWND 410
/ DB 352 LHYLKLIIKETLRLHPVAPLLVPRESTRDVIRGYHIPAKTRVFINAWAIGRDPKSWENA 411
/
/ QY 411 TCKPFRF-----DVGEEKKLVAFGMRACPGEPMAQSVSFTLLGLLQCFDQK 461
/

```

Db 412 EEFLPERFVNNSVDFKQDFQLIPFGAGRCPCGIAFGISSVEISLANILYWNWE 467

RESULT 7

US-09-126-420A-19

Sequence 19, Application US/09126420A

Patent No. 6176753

GENERAL INFORMATION:

APPLICANT: BATAARD, YANNICK

APPLICANT: ROBINEAU, TIBURCE

APPLICANT: DURST, FRANCIS

APPLICANT: WERCK-REICHAERT, DANIELE

APPLICANT: DIDIERJEAN, LUC

TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP7B1 FROM HELIANTHUS

TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN

TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL

TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S

FILE REFERENCE: 03715.0032

CURRENT APPLICATION NUMBER: US/09/126.420A

CURRENT FILING DATE: 1998-07-30

PRIOR APPLICATION NUMBER: 60/054,351

PRIOR FILING DATE: 1997-07-31

NUMBER OF SEQ ID NOS: 27

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 19

LENGTH: 508

TYPE: PRT

ORGANISM: Petunia x hybrida

US-09-126-420A-19

Query Match 28.2%; Score 731.5; DB 4; Length 508;

Best Local Similarity 31.4%; Pred. No. 2.1e-59;

Matches 159; Conservative 112; Mismatches 206; Indels 29; Gaps 6;

Qy 9 VLFVLFLGVKVFQSRKLRNIPGPPPLPIIGNLNLEQPIHRRFFORMSKQYGVNYSW 68

Db 12 LIFLTHIFITLSITNGRLPGPGWPGVIGALPILGAMPHVSLAKMAKYGAINYLK 71

Qy 69 FGSRLAVISSPTAYQBCFTKHDVALANRLPSLSKGKIFYNNTTVGSCSGHEHWRNLRRI 128

Db 72 VGTGCMVAVSTPDAKAFKLTDLNFSNRPNAGATHLAYGAQDMVFHYGPRWKLRLK 131

Qy 129 TALDVLSTQVHVSFGSIRSDTKLMQLVLAKNSEEPARVEISSMFDNLTNNMRM 188

Db 132 SNLHMGKALENANVRANLGHMLKSMFDMSEGE-----RVVVAEMLTFAMANTIGQV 187

Qy 189 ISGRKRYFGESEEMKNVBEAREFRFTVTETMLMGLMGLANKGDHLPFLRWFDONVEKRLSI 248

Db 188 ILSKRVF-----VKNKGVNEFKDMVVELMTTAGYFNIGDFIPCLAMNDLOGIEKMKRL 242

Qy 249 SKRYDSIILNKILHENRASNDQNSMIDHLLKLQETQPOYYTDQI-----IKGLALAMLFGG 304

Db 243 HKKFDALLTRMDFDEHKATSYERKGPDPFLDCVMENRDNSEGERLSTTNKALLNLEFTAG 302

Qy 305 TDSSTGTLEWLSNLNHPVLKARDELDTQVGDRLNEDSLPKLPYLKILLETLLRL 364

Db 303 TDTSSAIEWALAEWKNPALKKKAQGEQVIGNRRRLLESDIPNUPYRAICKETFRK 362

Qy 365 YPPAPILIPHVSSEDIITEGNIIPRDTIIVINGMQRDPOLMNDATCFKPERF-----418

Db 363 HPSTPLNLPRISENPICVDGYITPKNTLSNVAIGRDPVEMNPLEFYPERFLSGRNS 422

Qy 419 --DVEGEEKVLVAGMGRRACGFBPMQSVFTLGLLIOCFQDWKRVSEK-KLDMTENN 475

Db 423 KIDPRGNDFFELIPFGAGRRICAGTRMGIVMVEYILGILVHFSFKWKLPSVIELNMEBAFG 482

Qy 476 ILSRLPLEAMCKAR-----PLA 494

Db 483 LALQKAVEAMVTPRPIDVYAPLA 508

RESULT 8

US-08-606-505B-67

Sequence 67, Application US/08606505B

Patent No. 6114601

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro

APPLICANT: KIKUCHI, Shigeto

APPLICANT: SHIMADA, Yukihisa

APPLICANT: OHBAYASHI, Masaya

APPLICANT: SHIMADA, Ritsuko

APPLICANT: OKINAKA, Yasushi

TITLE OF INVENTION: NOVEL PLANT GENES

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10112-3801

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

COMPUTER: IBM PS/V

OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/606,505B

FILING DATE: 23-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP44963/92

FILING DATE: 02-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Perry, Lawrence S.

REGISTRATION NUMBER: 31865

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-218-2100

TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 67:

SEQUENCE CHARACTERISTICS:

LENGTH: 523 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Campanula medium

FEATURE:

NAME/KEY: CDS

LOCATION: 180 to 1748

IDENTIFICATION METHOD: by experiment

US-08-606-505B-67

Query Match 28.2%; Score 730; DB 3; Length 523;

Best Local Similarity 31.3%; Pred. No. 3.1e-59;

Matches 159; Conservative 117; Mismatches 208; Indels 24; Gaps 7;

Qy 4 VVSAYVLFVLFLGVKVFQSRKLRNIPGPPPLPIIGNLNLEQPIHRRFFORMSKQYGN 63

Db 12 LVAAISLIATYSFIRFLPFSHHHLPPOPTGWPILGALPLGTGPHVSLADMAVKYGP 71

Qy 64 VVSLWFGSRLAVISSPTAYQBCFTKHDVALANRLPSLSKGKIFYNNTTVGSCSGHEHWR 123

Db 72 IMYKLGSGKTVASNPKAARAFKTHDANFSNRPIDGGPTYLAYNAQDMVFYAEYGPWK 131

Qy 124 NLRRTALDVLSTQVHVSFGSIRSDTKLMQLVLAKNSEEPARVEISSMFDNLTNN 183

Db 132 LLKLGKLSHMLGPKALELWAKVSVGVHMLKEMY--EQSSKSVPPVPPVVPPELTAYMAN 189

Qy 184 NTRMISGRKRYFGESEEMKNVEEA-----REFRFTVTETMLMGLMGLANKGDHLPFLRWFDQ 239

Db 190 MIGRIILSRPFRPVITSKLDSSASASVSEFQVMVLMRMAGLFGIDFPIIWMMDLQ 249

Qy 240 NVEKRLKISIKRYDSIILNKILHEN-RASNDQNSMIDHLLKLQETQPOYYTDQI-----292

Db 250 GIORDMKVQKQKFDVLLNKNWIKETESADHRKDN--PDFLDILMAATQENTEGIQNLNLYN 307

```

Qy 293 IKGLALMFGTDSSTGTLWLSLNLNHNHPEVLKARDELDTQVGDRLLNESDLKPLP 352
Db 308 VKALLDLFTAGTDTSSVVIEWALAEMLNHRQILNRAHEENDQVIGRRRLQSDIENLP 367
Qy 353 YLRKIILTLRYPPAPILPHVSESDITIGFNIIPRTIIVINGWGMORPQWLNDATC 412
Db 368 YFOAICKETFRKHPSTPLNLPRISTEACEVDGFHPKNTRLIIVNIWAIGRDPKWNPLD 427
Qy 413 FKPERF-----DVEGEKKLVAFGMGRACGCEPMAMOSVFTLLGLLIQCFDWKRV - 463
Db 428 FTPERFLESEKHAIDPRGNHFLPFAGARRICAGARMGAASVEYLGLTLVHSPDWKLPD 487
Qy 464 SEEDLMTENNMTLSRLIPLEAMCKAR 491
Db 488 GVVEVNMEESFGIALQKKVPLSAIVTPR 515

RESULT 9
US-09-616-990-67
; Sequence 67, Application US/09616990
; Patent No. 6232109
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; KIKUCHI, Yasuhiro
; SHIMADA, Yukihisa
; OHBAYASHI, Masaya
; SHIMADA, Ritsuko
; OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
; COMPUTER: IBM PS/V
; OPERATING SYSTEM: MS-DOS Ver3.30
; SOFTWARE: PATENT AID Ver1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/616,990
; FILING DATE: 14-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Perry, Lawrence S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Campanula medium
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 180 to 1748
; IDENTIFICATION METHOD: by experiment
; SEQUENCE DESCRIPTION: SEQ ID NO: 67
US-09-616-990-67
Query Match 28.2%; Score 730; DB 3; Length 523;
Best Local Similarity 31.3%; Pred. No. 3.1e-59;

```

```

Matches 159; Conservative 117; Mismatches 208; Indels 24; Gaps 7;
Qy 4 VVSAYVLFLVFLGVKFVQSRKLRNIPPGPPPLPIIGNLNLEQIHRFFORMSKQYGN 63
Db 12 LVAAISLYATYSFIRFLPKPSHHHLPGPTGWPPIIGALPLLGTPHVSADMAVKYGP 71
Qy 64 VWSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLGSKYIFYNNTTVGSCSHEHWR 123
Db 72 IMYKLGSGTGVASNPKAARAFKTHDANFNSRPIDGGPTLYLAYNAQDMVFAEYGPWK 131
Qy 124 NRRITALDVLSTQRVHVSFGIRSDTKRLMQRLVLAKNSNEEFARVEISSMFNDLTYN 183
Db 132 LRKGLCSLHMLGPKALEDAHVVKVSVGHMLKEMY--EQQSKSVPVVVVPEMLTYMAN 189
Qy 184 NIMRMISGRFYGESEEMKNVEA----REFRTVTVMLEMLGLANKGDHLPFLRWFDFQ 239
Db 190 MIGRIILSRPFVITSKLDSSASASVSEFOYVMVLMEMAGLFNIGDFIPYIAWMDLQ 249
Qy 240 NVEKRLKSISRYDSIILKILHEN-RASNDQRQNSMIDHLKLOETQPOYYTDOI----- 292
Db 250 GIQDMKVIQKFDVLLNKMKEHTESAHDKDN--PDFLDILMAATQENTEGIQNLNVN 307
Qy 293 IKGLALMFGTDSSTGTLWLSLNLNHNHPEVLKARDELDTQVGDRLLNESDLKPLP 352
Db 308 VKALLDLFTAGTDTSSVVIEWALAEMLNHRQILNRAHEENDQVIGRRRLQSDIENLP 367
Qy 353 YLRKIILTLRYPPAPILPHVSESDITIGFNIIPRTIIVINGWGMORPQWLNDATC 412
Db 368 YFOAICKETFRKHPSTPLNLPRISTEACEVDGFHPKNTRLIIVNIWAIGRDPKWNPLD 427
Qy 413 FKPERF-----DVEGEKKLVAFGMGRACGCEPMAMOSVFTLLGLLIQCFDWKRV - 463
Db 428 FTPERFLESEKHAIDPRGNHFLPFAGARRICAGARMGAASVEYLGLTLVHSPDWKLPD 487
Qy 464 SEEDLMTENNMTLSRLIPLEAMCKAR 491
Db 488 GVVEVNMEESFGIALQKKVPLSAIVTPR 515

RESULT 10
US-08-948-564-4
; Sequence 4, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminsky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401

```

```
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-546-4

Query Match      28.1%; Score 729; DB 3; Length 510;
Best Local Similarity 31.1%; Pred. No. 3.7e-59;
Matches 158; Conservative 116; Mismatches 192; Indels 42; Gaps 10;

QY  7 YAVLFLVFLGVKVF-----QSRKLRNIPPGPPPLPIIGNLNL--EQPIHFFQ 55
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 FSIYFITSILFIFVFKLQVRSKTSSTCKLPGPRTPLIGNIHQIVGSLPVHYLK 70
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 56 RSKQYGNVSLWFGSLAVISPTAYQECFTKHDVALANRLPSLSGKVIFFYNNVTGVS 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 NLADKYGLMHLKLGESVNIIVTSPMAQEIIMTKDNLNFDSPDFVLSRVSYNGSGIVF 130
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 CSHEHWRNLRITALDVLSTQVHSPSGIRSDTKLMQRLVLAKNSNEEEARVEISS 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 SQHGDYWRQURKICTVELLTAKRVQSPRSIREEEVAELVKK--IAATASEGGSIFNLQ 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 MFNDLTNNIMRISGRFYGESESEMKNVBEAREFRETVTMELEMLGLANKGDHLPFLRW 235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 SIYSMTFG-----IAARAEGKKSRYQV-----FISNMHQLMLLGGFVADLYESSRV 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 FDFQNVKRLKISKRYDSILNKLHE-----NSRASNDRQ--NSMIDHLLKLOFTQPOYYT 289
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 FQMGATGKLEKVRHVTDRVLQDIIDEHKNRNSSEEREAEVDLDVLLKFQKSEFRLT 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 DOIKGLALAMLPGGTDSSGTLEWSLSNLLNHPVLKARDELDTQVGDRLNEDSLP 349
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 DDNIKAVIQDIFFGGTSSTSVWEGMSLIRPNVMEEAQAEVRRYVDSKGVYDETELH 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 KLPYLRKIILETLRLYPAPILIPHVSSDITIEGFNI PRDTVIINGMQRDPOLWND 409
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 QLIYLSKIIKETWRLHPVPLVPRVSRRCQINGEIPSKTRIIINAWAIGRPKYWGE 418
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 ATCFKPERF-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCDFWK--- 461
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 TESFKPERFLNSSIDRGTDFEFIFPGAGRRICPGITFAIPNTELPLAQLLYHFDWKLPN 478
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 462 RVSEKLDNTENNWITLSR-----LIPJ 484
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 KMKNEELDMTESGITLRRQNDLCLIPJ 506
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-08-532-065B-2
; Sequence 2, Application US/08532065B
; Patent No. 5753507
; GENERAL INFORMATION:
; APPLICANT: Ohta, Daisaku
; APPLICANT: Mizutani, Masaharu
; TITLE OF INVENTION: Plant Geraniol/Nerol 10-Hydroxylase and
; TITLE OF INVENTION: DNA Coding Therefor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 5753507artis Corporation
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: NJ
; COUNTRY: USA
; ZIP: 07936
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,065B
```

```
; FILING DATE: 22-SEP-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-532-065B-2

Query Match      28.0%; Score 725.5; DB 1; Length 495;
Best Local Similarity 34.6%; Pred. No. 7.4e-59;
Matches 165; Conservative 103; Mismatches 182; Indels 27; Gaps 9;

QY  1 LLAVVSYAVLFLVFLGVKVFQSRKLRNIPPGPPPLPIIGNLNLLEQPIHFFQFMSKQ 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  9 LLLLFCFLSCFLIPTTRSGRISRGATALPPGPPRLPIIGNIHLVGKHPHRSFAELSKT 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 YGNVSLWFGSLAVISPTAYQECFTKHDVALANRLPSLSGKVIFFYNNIT--VGSCSHG 119
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 YQVMSLKLGSNTVVIASPEAREVLRTHDQILSARSPNAVRSINHODASIVLWLPSSS 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 EHRNLRITALDVLSTQVHSPSGIRSDTKLMORLVLAKNSNEEEFARVEISSMFND 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 ARWRLRLSVTQLLSQRIEATKALRMNKVKELVS--FISESSDREE--SVDISRVAFI 184
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 LTYNNIMRISGRF--YGESESEMKNVBEAREFRETVTMELEMLGLANKGDHLPFLWPD 237
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 TTLNITSNLSFSDVLSGNAKASINGVQ-----DTVISVMDAAAGTPDAANTFFPFLRFLD 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 238 FQNVKRLKISKRYDSILN-----KILHENRASNDR---QNSMIDHLLKLOFTQPOYYT 289
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 LQGVKTFKVCYTERLVVRVFRGIDAKIAEKSSQNNPKDVSXKNDVNDLLDYKGESELSI 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 DOIKGLALAMLPGGTDSSGTLEWSLSNLLNHPVLKARDELDTQVGDRLNEDSLP 349
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 SD-IEHLLDMFTAGTDTSSTLEWPMTELLKNPKTKMAKAQAEIDCVIGQNGIVEESDIS 357
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 KLPYLRKIILETLRLYPAPILIPHVSSDITIEGFNI PRDTVIINGMQRDPOLWND 409
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 KLPYLOAVVYKETFRLHTPVLLIPRKAESDABILGFWMVKDTQVLVNWVAIGRDPVWDN 417
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 ATCFKPERF-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCDFWK 461
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 418 PSQFPERFLGKMDVGRDYEITPFGAGRRICPGMPLAMKTVSLMLASLYSFDWK 474

RESULT 12
US-09-126-420A-16
; Sequence 16, Application US/09126420A
; Patent No. 6376753
; GENERAL INFORMATION:
; APPLICANT: BATARD, YANNICK
; APPLICANT: ROBINEAU, TIBURCE
; APPLICANT: DURST, FRANCIS
; APPLICANT: WERCK-REICHT, DANIELE
; APPLICANT: DIDIERJEAN, LUC
; TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS
; TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN
; TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
; TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S
; FILE REFERENCE: 03715.0032
; CURRENT APPLICATION NUMBER: US/09/126,420A
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/054,351
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 27
```